

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 16:15:49 : Search time 172.18 Seconds
(without alignments)
472.225 Million cell updates/sec

Title: US-09-052-089a-2
Perfect score: 2393
Sequence: 1 MPILSLCTICSDPFHSDV.....VRIKTVSSASQPKDIFLCQ 470

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._virus:*
16: sp._bacteria:*
17: sp._archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	2315	96.7	470	11	008854
2	2306	96.4	470	11	09C6P4
3	1826.5	76.3	469	4	09BWF2
4	1798.5	75.2	469	4	000467
5	1045	43.7	233	11	0922M8
6	919	38.4	433	13	09YGN2
7	281.5	11.8	455	5	09V8D7
8	280.5	11.7	435	5	09S5S5
9	221	9.2	506	10	09M143
10	204.5	8.5	425	5	P90890
11	204.5	8.5	2473	11	090284
12	200.5	8.4	1690	5	044829
13	195.5	8.2	1690	5	09VTE5
14	191.5	8.0	2139	5	007569
15	190	7.9	2017	5	094992
16	190	7.9	2056	5	09W0W8

17	190	7.9	2057	5	094987	094987 drosophila
18	189.5	7.9	1447	11	090YF2	090YF2 mus musculus
19	189.5	7.9	1487	11	090YF3	090YF3 mus musculus
20	188.5	7.9	574	4	09H810	09H810 homo sapien
21	188.5	7.9	2138	5	09X2E3	09X2E3 amoeba prot
22	186.5	7.8	1156	17	028714	028714 archaeoglob
23	184	7.7	1941	5	026079	026079 placopecten
24	183.5	7.7	1937	13	091BD4	091BD4 gallus gall
25	183	7.6	1219	5	09N23	09N23 aequipecten
26	181.5	7.6	1871	5	09NCL3	09NCL3 drosophila
27	181	7.6	324	11	09EON8	09EON8 mus musculus
28	181	7.6	1156	16	066878	066878 aquifex aeo
29	181	7.6	1933	13	090337	090337 cyprinus ca
30	181	7.6	1940	5	09U7E3	09U7E3 pecten maxi
31	181	7.6	2007	13	002015	002015 gallus gall
32	180	7.5	692	5	096720	096720 dermatophag
33	180	7.5	1992	5	021440	021440 caenorhabd
34	179.5	7.5	1709	4	09UPS8	09UPS8 homo sapien
35	179.5	7.5	1710	4	09H1Q1	09H1Q1 homo sapien
36	179	7.5	1092	13	090338	090338 cyprinus ca
37	179	7.5	1950	5	026080	026080 placopecten
38	178.5	7.5	471	4	09BNG0	09BNG0 homo sapien
39	178.5	7.5	1999	11	063731	063731 rattus norv
40	178	7.4	482	5	09VD21	09VD21 drosophila
41	178	7.4	555	5	0961U3	0961U3 drosophila
42	178	7.4	1229	5	09N122	09N122 aequipecten
43	178	7.4	1243	5	09N121	09N121 aequipecten
44	178	7.4	1253	5	09N020	09N020 aequipecten
45	178	7.4	1790	3	007380	007380 saccharomyc

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	470 AA.
008854	AC	008854:		
008854	DT	01-JUL-1997 (TREMBLrel. 04, Created)		
008854	DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)		
008854	DE	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
008854	DE	MTrip.		
008854	GN	TRAIP OR MTRIP.		
008854	OS	Mus musculus (Mouse).		
008854	OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
008854	OC	Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.		
008854	OX	NCBI_TaxID=10090;		
008854	RP	[1]		
008854	RP	SEQUENCE FROM N.A.		
008854	RX	MEDLINE=97258620; PubMed=9104814;		
008854	RA	Lee S.Y., Lee S.Y., Choi Y.;		
008854	RT	"TRAF-Interacting protein (TRIP): a novel component of the tumor		
008854	RT	necrosis factor receptor (TNFR)- and CD30-TRAF signaling complexes		
008854	RT	that inhibits TRAF2-mediated NF-kappaB activation.";		
008854	RL	J. Exp. Med. 185:1275-1285(1997).		
008854	CC	-/- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.		
008854	DR	EMBL: U77844; AAB52994.1; -		
008854	DR	MGI: 1096377; Traip.		
008854	DR	InterPro: IPR001841; Znf_ring.		
008854	DR	Pfam: PF00097; zfc-C3HC4; 1.		
008854	DR	SMART: SM00184; RING; 1.		
008854	KW	Zinc-finger.		
008854	SO	SEQUENCE 470 AA; 53191 MW; 00FD705B52645133 CRC64;		

Query Match	96.7%;	Score 2315;	DB 11;	Length 470;
Best Local Similarity	97.7%;	Pred. No. 4e-116;		
Matches 459;	Conservative 0;	Mismatches 11;	Indels 0;	Gaps 0;
OY	1	MPILSLCTICSDPFHSDVAHICGHTFHLQCLIQWETAPSRCTCPOCRIOVGKTTIN 60		
Db	1	MPILSLCTICSDPFHSDVAHICGHTFHLQCLIQWETAPSRCTCPOCRIOVGKTTIN 60		

```

QY 61 KLFDLAQEEENVDAEFLKNELDVSKAQLSOKREKRDQAIIIDTLTLDTEENAVTES 120
    |||
Db 61 KLFDLAQEEENVDAEFLKNELDVSKAQLSOKREKRDQAIIIDTLTLDTEENAVTES 120
QY 121 LQNALNKAEMICSTLKQMKFLEORODETKOAREBAHRLCKMKMTMEQIEILLQSORSEV 180
    |||
Db 121 LQNALNKAEMICSTLKQMKFLEORODETKOAREBAHRLCKMKMTMEQIEILLQSORSEV 180
QY 181 EEMLRDMGVGSAVEQLAVYCVSLKKEYENLKEARKATGELADRLKDLVSSRSKLTN 240
    |||
Db 181 EEMLRDMGVGSAVEQLAVYCVSLKKEYENLKEARKATGELADRLKDLVSSRSKLTN 240
QY 241 TELDQAKLELSAQKDLSADQETISLRKSDPPGNLEPASATNETVSRVLFESPAPVE 300
    |||
Db 241 TELDQAKLELSAQKDLSADQETISLRKSDPPGNLEPASATNETVSRVLFESPAPVE 300
QY 301 MAMNRLHOPPPGDEIDLNTTFDVNTPTQTSGSQHCLPKKLCLEBRANSPMONVLKVKHV 360
    |||
Db 301 MAMNRLHOPPPGDEIDLNTTFDVNTPTQTSGSQHCLPKKLCLEBRANSPMONVLKVKHV 360
QY 361 SKPESQSLSGQRCVGEDELGAFLPFRNAVVGQKQPKRTAESRSTDVVRIGFDG 420
    |||
Db 361 SKPESQSLSGQRCVGEDELGAFLPFRNAVVGQKQPKRTAESRSTDVVRIGFDG 420
QY 421 LGGRTKFIOPRDTTIIIRPVVSKAKSKOKVRITVSSASQPKLDTELCQ 470
    |||
Db 421 LGGRTKFIOPRDTTIIIRPVVSKAKSKOKVRITVSSASQPKLDTELCQ 470

```

RESULT 2

Q9CNP4 PRELIMINARY; PRT; 470 AA.

```

AC 09CNP4:
DT 01-JUN-2001 (TREMBLERL 17, Created)
DT 01-JUN-2001 (TREMBLERL 17, Last sequence update)
DT 01-DEC-2001 (TREMBLERL 19, Last annotation update)
DE TRAF-INTERACTING PROTEIN.
GN TRAF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinaigawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Atakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kodota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welte C., Whitaker C., Wilming L.,
RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC - SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC EMBL: AK012948; BAB28567.1; -.
CC EMBL: AK012786; BAB28469.1; -.
CC MGD: MGI:1096377; Trajp.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF00097; zf-C3HC4; 1.
DR SMART: SM00184; RING; 1.

```

KW Zinc-finger.
SQ SEQUENCE 470 AA; 53149 MW; EBFARCA9A9F4BFZE CRC64;

Query Match 96.4%; Score 2306; DB 11; Length 470;
Best Local Similarity 97.2%; Pred. No. 1.2e-115;
Matches 457; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

```

QY 1 MPILSLCTICSDPFDSRDVAALHCGHTFHLQCLIQWETAPSRPCQCRIOVGKKTIN 60
    |||
Db 1 MPILSLCTICSDPFDSRDVAALHCGHTFHLQCLIQWETAPSRPCQCRIOVGKKTIN 60
QY 61 KLFDLAQEEENVDAEFLKNELDVSKAQLSOKREKRDQAIIIDTLTLDTEENAVTES 120
    |||
Db 61 KLFDLAQEEENVDAEFLKNELDVSKAQLSOKREKRDQAIIIDTLTLDTEENAVTES 120
QY 121 LQNALNKAEMICSTLKQMKFLEORODETKOAREBAHRLCKMKMTMEQIEILLQSORSEV 180
    |||
Db 121 LQNALNKAEMICSTLKQMKFLEORODETKOAREBAHRLCKMKMTMEQIEILLQSORSEV 180
QY 181 EEMLRDMGVGSAVEQLAVYCVSLKKEYENLKEARKATGELADRLKDLVSSRSKLTN 240
    |||
Db 181 EEMLRDMGVGSAVEQLAVYCVSLKKEYENLKEARKATGELADRLKDLVSSRSKLTN 240
QY 241 TELDQAKLELSAQKDLSADQETISLRKSDPPGNLEPASATNETVSRVLFESPAPVE 300
    |||
Db 241 TELDQAKLELSAQKDLSADQETISLRKSDPPGNLEPASATNETVSRVLFESPAPVE 300
QY 301 MAMNRLHOPPPGDEIDLNTTFDVNTPTQTSGSQHCLPKKLCLEBRANSPMONVLKVKHV 360
    |||
Db 301 MAMNRLHOPPPGDEIDLNTTFDVNTPTQTSGSQHCLPKKLCLEBRANSPMONVLKVKHV 360
QY 361 SKPESQSLSGQRCVGEDELGAFLPFRNAVVGQKQPKRTAESRSTDVVRIGFDG 420
    |||
Db 361 SKPESQSLSGQRCVGEDELGAFLPFRNAVVGQKQPKRTAESRSTDVVRIGFDG 420
QY 421 LGGRTKFIOPRDTTIIIRPVVSKAKSKOKVRITVSSASQPKLDTELCQ 470
    |||
Db 421 LGGRTKFIOPRDTTIIIRPVVSKAKSKOKVRITVSSASQPKLDTELCQ 470

```

RESULT 3

Q9BWF2 PRELIMINARY; PRT; 469 AA.

```

AC 09BWF2:
DT 01-JUN-2001 (TREMBLERL 17, Created)
DT 01-JUN-2001 (TREMBLERL 17, Last sequence update)
DT 01-DEC-2001 (TREMBLERL 19, Last annotation update)
DE TRAF INTERACTING PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG CARCINOMA;
RA Strausberg R.;
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC EMBL: BC000310; AAH00310.1; -.
CC DR InterPro: IPR001841; Znf_ring.
CC DR Pfam: PF00097; zf-C3HC4; 1.
CC DR SMART: SM00184; RING; 1.
CC KW Zinc-finger.
SQ SEQUENCE 469 AA; 53294 MW; B9EF3808FBC5985B CRC64;

```

Query Match 76.3%; Score 1826.5; DB 4; Length 469;
Best Local Similarity 77.6%; Pred. No. 4.3e-90;
Matches 363; Conservative 41; Mismatches 63; Indels 1; Gaps 1;
1 MPILSLCTICSDPFDSRDVAALHCGHTFHLQCLIQWETAPSRPCQCRIOVGKKTIN 60
|||

Db	1	MIRALCTICSPFDBSRVAAIHGHFTHLOCLLOMFETASRRCPCORIYQKRTIIN	60
Oy	61	KLEPDLAEEENVLDAEFLKNELDVKAOLSKDKREKRDQAIIIDLRTDLEERNATVES	120
Db	61	KLEPDLAEEENVLDAEFLKNELDVKAOLSKDKREKRDQAIIIDLRTDLEERNATVES	120
Oy	121	LONLNKAEMLCSTLKQKMKFLEORODEFKQAKREEHNRKCKMKMTMEQTELLLOSRSV	180
Db	121	LONLNKAEMLCSTLKQKMKFLEORODEFKQAKREEHNRKCKMKMTMEQTELLLOSRSV	180
Oy	181	EEMIDMGVGSAYBOLAVYCVSLKREYENLKARKATGELADRIKDLVSSRSKLTLN	240
Db	181	EEMIDMGVGSAYBOLAVYCVSLKREYENLKARKATGELADRIKDLVSSRSKLTLN	240
Oy	241	TLLDQAKLELRQAQNDLQSADEITSLKSSDPNGNLEPASATNVTYSRLVFESAPVE	300
Db	241	TLLDQAKLELRQAQNDLQSADEITSLKSSDPNGNLEPASATNVTYSRLVFESAPVE	300
Oy	301	MNPNLHOPPEDELDLNTTFDVNTPPQOTSGSOCLPKLCLLEASRSPMWNLVKHVY	360
Db	301	MNPNLHOPPEDELDLNTTFDVNTPPQOTSGSOCLPKLCLLEASRSPMWNLVKHVY	360
Oy	361	SKPESQLSLGGORCYGELDELGAAPLEFIRNAVLGQOKPNFTAESRSTDVYRIGEDG	420
Db	360	PKRESQSLSGGSCAGEPDEELVGAFFIVRRAIIGOKPKPRRESSCSKDVYRTGFDG	419
Oy	421	LGGRKFTIOPRDTIIRPVVSKSKSKOVKITVYSSASQPKLDFEL	468
Db	420	LGGRKFTIOPRDTIIRPVVSKSKSKOVKITVYSSASQPKLDFEL	467
RESULT	4		
000467			
000467	PRELIMINARY:	PRT:	469 AA.
AC	000467:		
DT	01-JUL-1997 (TREMBLrel. 04, Created)		
DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	HTRP.		
GN	HTRP.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=97258620; PubMed=9104814;		
RA	Lee S.Y., Lee S.Y., Choi Y.;		
RT	"TRAF-interacting protein (TRIP): a novel component of the tumor		
RT	necrosis factor receptor (TNFR)- and CD30-TRAF signaling complexes		
RL	that inhibits TRAF2-mediated NF-kappaB activation.";		
CC	J. Exp. Med. 185:1275-1285(1997).		
CC	-1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.		
DR	EMBL: U77845; AAB52993.1; -		
DR	InterPro: IPR001841; Znf_ring.		
DR	Pfam: PF00097; zf-C3HC4.1.		
DR	SMART, SM00184; RING.1.		
KM	Zinc-finger.		
SEQUENCE	469 AA; 53138 MW; 2054ED04B84ABAE4 CRC64;		

Query Match	75.2%;	Score 1798.5;	DB 4;	Length 469;
Best Local Similarity	76.9%;	Pred. No. 1.3e-88;		
Matches 360;	Conservative 41;	Mismatches 66;	Indels 1;	Gaps 1;

QY	1	MPILSLCTICSDPFHSDVAHICGHTFHLQCLIQMFEPASPTGQCSQIYQKRTIN	60
Db	1	MPIRALCTICSDPFHSDVAHICGHTFHLQCLIQSEFASPTGQCSQIYQKRTIN	60
QY	61	KLFEDLAEENVDLDAEFLKNEELDSVAAQASQKDKREKRDSQAIIIDLRTLLEENATVES	120
Db	61	KLFEDLAEENVDLREFLNKELNENVAQASQKDKREKDSQIIDLRTLLEENATVES	120

Qy	121	LONLNKAEMLSTLTKKQMKLEORODETKOAREBNHRLCKKMTMQIELLQSORSEV	180
Qy	121	LONLNKAEMLSTLTKKQMKLEORODETKOAREBNHRLCKKMTMQIELLQSORSEV	180
Db	121	LQOALGKAEMLSTLTKKQMKLEOODETKOAEGRALRSKMTMQIELLQSORSEV	180
Qy	181	BEHIMDMCVGSAAVQQLAVVCSVLKKEVENLKKARKTKTGELADLKKDYLSRSSKLTIN	240
Db	181	BEHIMDMCVGSAAVQQLAVVCSVLKKEVENLKKARKTKTGELADLKKDYLSRSSKLTIN	240
Qy	241	TELDQAKLELSAQKDLOSAOQETLSLTKSDDBPGNLEPASATNEVSRVFEPSAPVE	300
Db	241	TELDQAKLELSAQKDLOSAOQETLSLTKSDDBPGNLEPASATNEVSRVFEPSAPVE	300
Qy	301	MNPNRLHPPEDELDLTTEDVTNPPROTGSQONCLPKLCLERARSPMQLNKKVY	360
Db	301	MNPNRLHPPEDELDLTTEDVTNPPROTGSQONCLPKLCLERARSPMQLNKKVY	360
Qy	361	SKPEBSOLSGORCVGELDEELAGAFPLEIRINAVLGOKOPNRHTAESRSSTDVVRIGEDG	420
Db	361	SKPEBSOLSGORCVGELDEELAGAFPLEIRINAVLGOKOPNRHTAESRSSTDVVRIGEDG	420
Qy	421	LGGRKFTQPDRTTIIIRVPVPKSKAKSKQKVRITVSSASDPKIDTFL	468
Db	421	LGGRKFTQPDRTTIIIRVPVPKSKAKSKQKVRITVSSASDPKIDTFL	468
Qy	468	LGGRKFTQPDRTTIIIRVPVPKSKAKSKQKVRITVSSASDPKIDTFL	468
Db	468	LGGRKFTQPDRTTIIIRVPVPKSKAKSKQKVRITVSSASDPKIDTFL	468

RESULT	5
Q922M8	
ID	Q922M8
PRELIMINARY;	
PRT;	223 AA.

DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)
DE TRAF-INTERACTING PROTEIN.

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006929; AAH06929.1; -.
SQ SEQUENCE 223 AA; 25584 MW; 412291C205C43071 CRC64;

Query Match	43.7%	Score 1045;	DB 11;	Length 223;
Best Local Similarity	99.08;	Pred. No. 8.6e-49;		
Matches 204; Conservative	1;	Mismatches 1;	Indels 0;	Gaps

Db 1 MPRALCTICSDFEHSDRVAALHCGHTFHLQCLIQFETAPSRPCPQCRIVGKTIIN 60

Db 61 KLFPLDAGEEAVLDAEFLKNELDVSVAQLSQKREKROQAIIDTLRDTLEERNATVES 120

06 181 EEMIRDMGOSAVEQILAVYCSLKK 206
07 |||||

RESULT 6
Q9YGN2

AC	01-MAY-1999 (TREMBLrel. 10, Created)
DT	01-MAY-1999 (TREMBLrel. 10, last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)


```

Db 183 ----EADPHVLAMVSTLKLRELROCESKKTLELRNVKVVQNDLREIELEKRAYHNIPNV 238
QY 242 -----ELDOAKLELBSA--QKDIASADOETSLRKS---DDPGNLEPASANEVSR 290
Db 239 FIIDMCSFDSCKLEERVSHLESPLYAOEKLQAEKNTAYLDSP----NMSCGLNSILA 294
QY 291 LVFE-----SPAVEMMNPRLHQPFGDEIDLNTTFDVNTPPTQSGSQHCLPKK--LC 342
Db 295 LKREERTTISPYKENIK-----RREESTSPYLN-KSSVGLAHLTLTKGNIG 343
QY 343 LERAR-SPMONLKKVHKVSKESPQSLGQRCVGELEDELAGAPLEFIRNAVLGOKOPN 401
Db 344 LAKSKIPIKGVGGVGMTS-----GTRKTSDDLSEKYSIF-----KKPR 384
QY 402 RTTAESSRSDVVRIG----FDGLGRTKFTQPRDPTIIRPVYKSAKSKOKVRITVS 457
Db 385 LLGSSSSSALTATTGTFNFMGMGSEK-----VDPFAQRAEEGLSTIRSSAL 435
QY 458 S 458
Db 436 S 436

RESULT 8
Q95SS5 PRELIMINARY; PRT; 435 AA.
ID Q95SS5
AC Q95SS5
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DE GH03577P.
GN GS140.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Garlin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celinker S.;
RU Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY060610; AL28158.1; -.
SQ SEQUENCE 435 AA; 48455 MW; 0BFF0AC81DD9416E CRC64;

Query Match 11.7%; Score 280.5; DB 5; Length 435;
Best Local Similarity 22.0%; Pred. No. 9.3e-08;
Matches 104; Conservative 99; Mismatches 187; Indels 83; Gaps 17;

QY 7 CTTCSDFFDSRDVAALHCGHTFHLOCLIQWFFETAPSTCPQCRIOYQKTIINKLEFDL 66
Db 6 CVICALEFGADVEFAVTCGMEFHNLNQLNQLDR--SKTCQCRNKCTTRNIF-RVFNFL 62
QY 67 AQEEENVLDEFLKNEIDSVYKAQLSOKDREKRSQAIITDLRPLEERNATVESIQNALN 126
Db 63 ANLDVSHIDVSGIQEOLDNAML SMKVKYERKNKDEQOIRDLKETQKCLKITAGLEQVQ 122
QY 127 KAEMLCTLKKOMKFLQRODETKQAREEAHRLCKKAKTMOIELLOSOSEVEEMIRD 186
Db 123 KKDPLISSYVEQIGVLKSDAHVGLKREKNTKLSQIOSMGISAILAAGSADADRLKN 182
QY 187 MGVSQSAVEQIAYCVSLKKEYENIK---EARKATGELADRLKKDIYSSR---SKLTL 239
Db 183 ----EADPHVLAMVSTLKLRELROCESKKTLELRNVKVVQNDLREIELEKRAYHNIPNV 238
QY 240 NTELDQAK-----LELSAQKDLQSDOETIS--LRKSSDDPGNLEPASATNETVSR 292
Db 239 ESDLYQAOEKLOAEKNTAYLDSPNASCGLNSILALKREERTTISP--TVKENIKR-I 295

```

```

QY 293 FESPAPVEMMNPRLHQPFGDEIDLNTTFDVNTPPTQSGSQHCLPKK--LCLERAR-SP 349
Db 296 EESTSPY-----LNT-----KSSVGLAHLTLTKGNIGLAKSKISP 331
QY 350 MONVLYKHHVSKESPQSLGQRCVGELEDELAGAPLEFIRNAVLGOKOPNRTTASRS 409
Db 332 IKGVGGVGMTS-----GTIRKTSDDLSEKYSIF-----KKPRLLGSSSS 372
QY 410 STDVVRIG----FDGLGRTKFTQPRDPTIIRPVYKSAKSKOKVRITVS 458
Db 373 SALTATGSMNFVYMGMGSEK-----VDPFAQRAEEGLSTIRSSALS 416

RESULT 9
Q9M143 PRELIMINARY; PRT; 506 AA.
ID Q9M143
AC Q9M143
DT 01-OCT-2000 (TEMBLrel. 15, Created)
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE PUTATIVE RING ZINC FINGER PROTEIN.
GN AT4G01270.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamer B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RU Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RU Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL: AL161491; CAB80936.1; -.
DR InterPro: IPR001841; Znf_fing.
DR Pfam: PF00097; zf-C3HC4; 1.
DR SMART: SM00184; RING; 1.
KW Zinc-finger.
SQ SEQUENCE 506 AA; 56518 MW; 1B06751A4E299F8 CRC64;

Query Match 9.2%; Score 221; DB 10; Length 506;
Best Local Similarity 21.9%; Pred. No. 0.00016;
Matches 120; Conservative 79; Mismatches 193; Indels 156; Gaps 24;

QY 5 SLCTICSD----FFDHSRDVAALHCGHTFHLOCLIQWFFETAPST---RTCPQCRIOYQKKT 57
Db 10 AICSDICYEDLKPVENLQSIISA--CGHVHFHLCIQWFFECPSYNNKRCPCCKOKCSIKD 67
QY 58 IINKLFPDLAQEEENVL-----DAEFLKNEIDSVYKAQLSOKDREKRSQAI 104
Db 68 PCRLYFOSSGQNTDSIDSKVYGTEEDPVYLKGEYKRLKGVQVLTSLAKKKKEENVEVS 127
QY 105 DTLRDLE-----ERNATVESIQNALNKAEMLCSTLKKOMKFLQRODETKQARE 154
Db 128 DKLHQCHDQLEKDKVKKRMEALQETSTQHLTKLASEBQIQNSQCVKLQER--TVLAL 184
QY 155 EAHRLKC-----KMKTYEQIELLOS---QRESEVEEMIR--DMGV 189
Db 185 ELASIKLVSDLSLEDDVLKALLGNNAKTKDTIDTLVKSIVIRNRSYKELACNOLGR 244
QY 190 GOS--AVQOLAVYCVSLKKEYENLKEARKATGELADRLKKDIYSSRSLKTLNTELDQAKL 248
Db 245 GEASSSERLEKALEKIEKRLKRMLELITTEERENRALROINVK---KCSYTYSEPAI 301
QY 249 ELRSQKDLQSDOETISLRKSSDDPGNLEPASATNETVSRVLESPAPVEMMNPRLHQ 308
Db 302 ESMSSFRLL-SSDNKV-----EKISTPQKLEKDGFTIQGSL-----RGRE 343

```

```

OY 309 PRFGDEIDLNTPTPVN---TPPTOTSG-----SOHCPLKCLERANSPMQNVL 354
Db 344 DSEVSRIT--SVIVDDDYVETNTSGIRDMNTNIEEKGDNSMVKDKIFNIRDPITSSV- 400
OY 355 KKVHVSKPEEO-----LSIGGRCVG-----ELDELAGAPLEIRNAVIGQK 399
Db 401 -----SPYNGSGNIMWSSGTRNRNLGRMSKHGERNEATPISLGSP----- 441
OY 400 PNRTAESRSSTDVYRIGFDGLGRTKFTIOPRDTTIRPVVSKAKSKO-----KVRI 453
Db 442 -----RRKDDILSIGDPDGKGRIK-----VLRSKPOISKTWNASSGSGRFRKLTG 484
OY 454 KTVSSASQ 461
Db 485 KTVSGSSQ 492

RESULT 10
P90990
ID P90990 PRELIMINARY; PRT; 425 AA.
AC P90990.
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE B0432.9 PROTEIN.
GN B0432.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;
OC Rhabdilitidae; Peloderiinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RC MEDLINE=94150718; PubMed=7906398;
RX Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Crexton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten U., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopre A., Saunders D., Showkneen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterson R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RT Nature 368:32-38(1994).
RN [2]
RP RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RC STRAIN-BRISTOL N2;
RC Waterston R.;
RA Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; U08036; AAB37893.1; -.
DR InterPro; IPR001841; Znf_fing.
DR Pfam; PF00097; zfc3HC4; 1.
DR SMART; SM00184; RING; 1.
DR ZINC-finger.
SQ SEQUENCE 425 AA; 48104 MW; 375F524E5C092B2D CRC64;

```

Query Match	8.5%	Score 204.5;	DB 5;	Length 425;
Best Local Similarity	25.98;	Pred. No. 0.001;		
Matches	73;	Conservative 49;	Mismatches 105;	Indels 55;
				Gaps 12.

Db	184	PTSLIOGSCSICFEDLKNOKNDIKSIAVCHYIHHCISOMIAT--KRCPCRCRTVRKNGF	24.1
Oy	59	INKLFPLDAAQ---EEENVLDAEFLAKNELDSYKAAOLSOKDREKDSQAIIDTLRDTLEERN	115
Db	242	VEKLEFEDVORNGGEAEKRPPELDYREEHY---KLSYSLVEODEKICLTATLTKNKLD--	294
Oy	116	ATVESLQNAL---NKAEMLCSYLLKOKMFLERODHETKOAREEABRLKCKMKTMEQIEL	171
Db	295	---TVKSLEKKIIRKEDKYKOEIPIKLOATINHLTISSEETAYLKRLOESKARLKTCEYKI	353
Oy	172	LLOOSRSEVEEMIRDMGVGSAYBOLAVYCSVLK---EVENLKEARKATGLADRLK	227
Db	354	LTVH-----SSADNDQJEGY---LKKNGINDTEKFFQMLKST-----NK	389
Oy	228	DVLSRSKLLKTLNTELDOAKLELRSAAOKDLOSADOETISLTK	269
Db	300	DLTKDR---REAAKEIDOLKKEVOS---LKRRAQOEONAAIK	424

RESULT	11		
Q9QZ84			
ID	Q9QZ84	PRELIMINARY:	PRT: 2473 AA.
AC	Q9QZ84;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)		
DE	LEK1 (FRAGMENT).		
GN	6530404A22RIK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
OX	NCHI_TaxID-10090;		
	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEEDLINE-99303627; PubMed-10373470;		
RA	Goodwin R.L., Pabon-Pena L.M., Foster G.C., Bader D.;		
RT	"The cloning and analysis of LEK1 identifies variations in the		
RL	LEK/centromere protein F/mitosis gene family.";		
DR	J. Biol. Chem. 274:18597-18604 (1999).		
DR	EMBL: AF194970; AAF07196.1		
MD	MG1:1915046; 6530404A22RIK.		
FT	NON_TER	1	
SEQ	NON_TER	1	
	2473 AA;	2481903 MW;	504E91A1A150A3B5 CRC64;

Query Match	8.5%;	Score 204.5;	DB 11;	Length 2473;
Best Local Similarity	20.5%;	Pred. No. 0.0063;		
Matches	95;	Conservative 101;	Mismatches 163;	Indels 105;
				Gaps 14;

Qy	56	KTNNKLEPPLAOEEN----	VIDAEFLNLEDSVYAOLOSQKDRKRSOAMIDIRPTL	111
Db	1629	KOTVETLERELERKEEENQOELADISENLAABEYTTLKAQDENTKSLRIPELDVTVKRER	1668	
Qy	112	EERNATVESLONALNKAMCSTLTKOMKFLDOR-----	ODETQOAREEA	156
Db	1689	ENLAKOLOEKOSRSELDEKSSLRLEIRLEEKQDARVOMEDSKSAMILOMQLKELREAV	1748	
Qy	157	HLCSKKNYTOEITELLLOSRSVEEMINDMGVGSANVQOLAVCVSLKKEYLENKARK	216	
Db	1749	AALCNDDETILKAQOSLIDOPGEEVHNH-----	SSITKRLVNHIDAOBKKNQNTLEOK	1801
Qy	217	ATVGLADRLK-----	KDLV-----SSRSKLTNTLEDAQKLEIRSAQOKL-	257
Db	1802	ESKHADRLKDVENLEQOELLSEKMNIMFOAEKSKAEIOTLSESTIORMAONQLODETLEI	1861	
Qy	258	-----OGADDEISLKKSDDPNGCINLEPRASAVNENYSR-----	LVFESPAVEMINPRH	307
Db	1862	STRSEENELIKELKEDEORSVSDLETTNPSTENILKQKDEOKOMKEAKITVEMLOQTOL	1921	
Qy	308	QRPGEDEIDINTTFVDVNTPTPTQSGSOHCLPKKICLERASBPQVNLKVKHVSKEPSQOL	367	
Db	1922	-----EINENYV-----	VSICNDQ-----EVSKTRKO-1942	

Db 724 QIOLEKESIEQOLAL-----KONELE-DFOKQSESEVHLOEIKAKONTQKDFELVESGES 777
Qy 104 IDTLRDTLEERNATIVESIONALNKAEMLCSTLKQMK-FLEORODETKOAREAHRLCK 162
Db 778 LKKLOOQLEQKTLGHEKIQAALEF-----LKKKEKTLIKKEQELQOLOSKESES 830
Qy 163 MKTME-QIEILLQSORSEVEEMIRDMGVGQSAVEQLAVYC---VSLKKEYNL-----KE 213
Db 831 LKVVQVQLEQLOQQAASAGEEGSKTVAKLHDEISQLKSOAERTQSELKSTQSLNLEAKSKQ 890
Qy 214 ARKAGELADRLKRD--LVSSRSKLTUNTELDOAKLELRSKQKLOKQSAODEITSLRKKS 271
Db 891 LEAAGSELEEKAKKSGHLEQITTKL-----SEVGERQAAALSCHTQVESKTKQ----- 939
Qy 272 DDPGNLEPASATNETVSRVLEFESPA-----PYEMNPNRLHQ----- 308
Db 940 -----LEAANAALAEKVAKKEVAESRAEASDDODKRYKETITDTLHALQAEKSSSALHLTKL 993
Qy 309 PPFGEI-----DLNTFFDVNTPTPTQTSQSOHCLPKLCLERARSPMQ----- 351
Db 994 SKFSEIATGHKELTSKAD-----AMSQEMLOKEKELEQLROQLDSDQSDQTKLKA 1045
Qy 352 -----NYLKKVHKVSKPESQSLSGQRCVGELEDELAGAPFLIRNAVIGOK 398
Db 1046 GERKEKSEESIKNLQEEVTKAKTENLELSTGTQTTIKDLERLE-----ITNLELOHK 1099
Qy 399 QPNRTTAESSRSSTDVYRIGFDGLGRTKFIQPRD-TTIIRPVYPSKAKSKOKVRIKTVS 457
Db 1100 -----EKMASEDAKQIA-----DKTIVEALQVANNANISATNLELSTVL 1138
Qy 458 SASQ 461
Db 1139 EVLQ 1142

RESULT 14
Q07369 PRELIMINARY; PRT; 2139 AA.
ID Q07569
AC Q07569; 002504;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MYOSIN HEAVY CHAIN.
GN MICA.
OS Entamoeba histolytica.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=5759;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HMI:IMSS;
RX MEDLINE=93295430; PubMed=8515774;
RA Raymond-Denise A., Sansonetti P., Guillen N.;
RT "Identification and characterization of a myosin heavy chain gene
(mhca) from the human parasitic pathogen Entamoeba histolytica."
RL Mol. Biochem. Parasitol. 59:123-131(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HMI:IMSS;
RA Guillen N.;
SUBMITTED (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; L03534; AAB46065.1; -.
DR HSSP; P08799; 1MND.
DR InterPro; IPR000048; IQ.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00242; MYSC; 1.
SEQUENCE 2139 AA; 245225 MW; C68307341DB51DD1 CRC64;

Query Match

8.0%; Score 191.5; DB 5; Length 2139;

Best Local similarity 22.1%; Pred. No. 0.027;
Matches 97; Conservative 83; Mismatches 161; Indels 97; Gaps 16;
Qy 66 LAQEEENVLDAEFLKNELDVSKAQLSQK-----DREKRSQALIIDTLRDLERNNTV 118
Db 1671 VAOEEEEK-----ORLESDIAELKEQLEOERTTAANAEAERKKIQAELEDEVKFNLEDTNQR 1726
Qy 119 ESL--QNALNKAEMLCSTLKQMKFLEORODETKOAREAHRLCKMKMTQIEILLQSQ 176
Db 1727 EKLVAKNSENDAEI--DSLKEKKALE--DEIEKTTDDNNKLS-----EIDSIDLRKY 1775
Qy 177 RSEVEEMIRDMGVGQSAVEQLAVYCVSLKKEYENLKEARKATGELADRLKRDVSSRSKL 236
Db 1776 NALLDSKSDVSMKKEKFPDELKVTKDIALETEKKNHAEFTMRILKG-----FLKEEAFAVQVRL 1831
Qy 237 KTLNTELDOAKLELRSKQKLOKQSAODEITSLRKSDDPGNLEPASATNETVSRVLEFESP 296
Db 1832 EALQKNLDAQOEKAKKRVRAADGELKSLNNLELDVKKDLDKRA----- 1876
Qy 297 APVEMNPNRLHQPPGEIDINTTTPD-----VNTPTQTSQSOHCLPKLCLERARSPMQ----- 345
Db 1877 -----QDDLADKEDELATLDOCKKTLVKKQSFDSRIQE-MQEQDLEKAGRAK 1924
Qy 346 ARSPQNVLKKVHKVSKPESQSL-----SLGQRCVGELEDELAGAPFLIRNA 393
Db 1925 AOKQKQAEYKLOELQENDNPFEYKETAADRINTLSAQ--DDLQKEL-----EK 1973
Qy 394 VLQKQPNRTTAE-----SRSSDVYRIGFDGLGRTKFIQPRD-TTIIRPVYPSKAKSK 447
Db 1974 ERGLKQDSEKEVQRLRVKCOELTKVAEVG--GANVSIKAKVAKYAEIEELTTEADAL 2031
Qy 448 KQVRIKTVSSASQPKLD 465
Db 2032 KAKMAEKKAKTSQKKLD 2049

RESULT 15
Q04992 PRELIMINARY; PRT; 2017 AA.
ID Q04992
AC Q04992; 024138;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NONMUSCLE MYOSIN-II HEAVY CHAIN.
GN ZIP OR CG15792.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Ephydroidea; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96144835; PubMed=8568878;
RA Mansfield S.G., al-Shirawi D.Y., Ketchum A.S., Newbern E.C.,
RA Kiehart D.P.;
RT "Molecular organization and alternative splicing in zipper, the gene
that encodes the Drosophila non-muscle myosin II heavy chain."
RL J. Mol. Biol. 255:98-109(1996).
DR EMBL; U35816; AAB09048.1; -.
DR EMBL; U35816; AAB09051.1; -.
DR HSSP; P10587; 1BR2.
DR FlyBase; FBgn0005634; zip.
DR InterPro; IPR000048; IQ.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.

DR SMART: SM00242; MYSC: 1.
DR PROSITE: PSS0096; IQ: 1.
SQ SEQUENCE 2017 AA; 232089 MW; 8C76FEF2EBD02EBE CRC64;

Query Match 7.9%; Score 190; DB 5; Length 2017;

Best Local Similarity 23.8%; Pred. No. 0.03; Mismatches 145; Indels 76; Gaps 14;
Matches 88; Conservative 61;

QY 56 KTITNKLPFDLQAEENVLDAE--FLKNELDSVQAQLSQKREKRDSDQ-----AI 103
D 1263 KTVLEK-----AKGTLEENADLATELRSVSSRQENRRKQAESQIAELQVKLAE 1314
QY 104 IDTLRDTLEER-----NATVESLONALNKAEMLCSTLKKQMKFLEBRODETKQAREEAHR- 158
D 1315 IERRSELQEKCTKLQOEAENITNQLAEALKASAAVKSASNMESQLTEAQOLLEETRQ 1374
QY 159 ---LCKMKMTMEQIELLQSQRSEVEEMIRDMGVQSAVEQLAVYCVSLKREYNLKEAR 215
D 1375 KLGLSSKLRQIYESKEKALQOELDEDEKRNYY---ERKLAVTTQOMQEIKKKAEDADLA 1431
QY 216 KATGELADRLKDLVSSRSKIKTL---NTELDQAKLELRSAQKQ---LQSADEITSLR 268
D 1432 KELEGGKKRLKKDIEALERQVKELIAQNDRLDKSKKIQSELEDATTELEAQRTKVELE 1491
QY 269 KKSDDPGNLEPASATNVTGRVLFESPAPVEMNPRHQPFQDEIDLNTTFPVNTPPT 328
D 1492 KKQK---NFKITLAEKAISEQIAQERDTAREAREKETVLSVSRLEDAFP----- 1541
QY 329 QTSQSQHCLPKKCLERASPMQNVL-----KKVHKVSKP---ESQLS-LGQ 372
D 1542 -----KIEDLENKRKTLQNELDLANTQGTADKNVHELKAKALQSALAEKQ 1591
QY 373 RCVGELDEEL 382
D 1592 N-BELEDL 1599

RESULT 16
Q9W0W8 PRELIMINARY; PRT: 2056 AA.
AC Q9W0W8;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE ZIP PROTEIN.
GN ZIP OR CG15792.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pharyngota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidae; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherter S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abtill J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
RA Butlis K.C., Butsam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunham B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegian C.,
RA Jatali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Mishna N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclik J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "the genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL; AE003465; AAF47311.1; -;
DR HSP; P10587; IBR2.
DR Flybase; FBgn0005634; zip.
DR InterPro; IPR000048; IQ.
DR InterPro; IPR001609; myosin_head.
DR InterPro; IPR004009; myosin_N.
DR InterPro; IPR002928; myosin_tail.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin_head; 2.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PSS0096; IQ; 1.
SQ SEQUENCE 2056 AA; 236627 MW; 1D074D9CEB1538E4 CRC64;

Query Match 7.9%; Score 190; DB 5; Length 2056;

Best Local Similarity 23.8%; Pred. No. 0.031; Mismatches 145; Indels 76; Gaps 14;
Matches 88; Conservative 61;

QY 56 KTITNKLPFDLQAEENVLDAE--FLKNELDSVQAQLSQKREKRDSDQ-----AI 103
D 1302 KTVLEK-----AKGTLEENADLATELRSVSSRQENRRKQAESQIAELQVKLAE 1353
QY 104 IDTLRDTLEER-----NATVESLONALNKAEMLCSTLKKQMKFLEBRODETKQAREEAHR- 158
D 1354 IERRSELQEKCTKLQOEAENITNQLAEALKASAAVKSASNMESQLTEAQOLLEETRQ 1413
QY 159 ---LCKMKMTMEQIELLQSQRSEVEEMIRDMGVQSAVEQLAVYCVSLKREYNLKEAR 215
D 1414 KLGLSSKLRQIYESKEKALQOELDEDEKRNYY---ERKLAVTTQOMQEIKKKAEDADLA 1470
QY 216 KATGELADRLKDLVSSRSKIKTL---NTELDQAKLELRSAQKQ---LQSADEITSLR 268
D 1471 KELEGGKKRLKKDIEALERQVKELIAQNDRLDKSKKIQSELEDATTELEAQRTKVELE 1530
QY 269 KKSDDPGNLEPASATNVTGRVLFESPAPVEMNPRHQPFQDEIDLNTTFPVNTPPT 328
D 1531 KKQK---NFKITLAEKAISEQIAQERDTAREAREKETVLSVSRLEDAFP----- 1580
QY 329 QTSQSQHCLPKKCLERASPMQNVL-----KKVHKVSKP---ESQLS-LGQ 372
D 1581 -----KIEDLENKRKTLQNELDLANTQGTADKNVHELKAKALQSALAEKQ 1630
QY 373 RCVGELDEEL 382
D 1631 N-BELEDL 1638

```
RESULT 17
Q94987 PRELIMINARY: PRT: 2057 AA.
ID Q94987
AC Q94987:
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NONMUSCLE MYOSIN-II HEAVY CHAIN.
GN ZIP OR CG15792.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96144835; PubMed=8568878;
RA Mansfield S.G., al-Shirawi D.T., Ketchum A.S., Newbern E.C.,
RA Kiehart D.P.;
RT "Molecular organization and alternative splicing in zipper, the gene
RT that encodes the Drosophila non-muscle myosin II heavy chain.";
RL J. Mol. Biol. 255:98-109(1996).
DR EMBL: U35816; AAB09049.1; -;
DR EMBL: U35816; AAB09050.1; -;
DR HSSP: P10587; 1BR2.
DR Flybase: FBgn0005634; zip.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR001609; myosin_head.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00663; myosin_head; 2.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
SQ SEQUENCE 2057 AA; 236642 MW; 0D59519C1E83A411 CRC64;
```

```
Query Match 7.9%; Score 190; DB 5; Length 2057;
Best Local Similarity 23.8%; Pred. No. 0.031;
Matches 88; Conservative 61; Mismatches 145; Indels 76; Gaps 14;

QY 56 KTIINKLFFDLQAEENVLDAE--FLKNELDYKQAQLSQKREKDSQ-----AI 103
DB 1303 KTVLEK-----AKGTLEAMNADLATELRVNSRQENDRRRRKQAESQIAELQVLA 1354
QY 104 IDTLRDLTLEER-----NATESIQNALNKAEMLCSTLKQMKFLERODETQOAREEHR- 158
DB 1355 IERRARSELQEKCTKIQQAENITNLEAEELKASAAVKSASMEQSLTEAQLLEETRRQ 1414
QY 159 ---LCKNKKTMQIELLLOSQRSEVEEMIRDMGVGQSAVEQLAVYCVSLKREYENLKEAR 215
DB 1415 KLGSSTKRLQRIESEKEALQOELEEDDEAKRMY---ERKLAETVTQMGIKKAEEDADLA 1471
QY 216 KATGELADRLKDKDVSSSKTKTL---NTELDQAKLELRSAQKQD-----LQSADEQITSLR 268
DB 1472 KELEBGKRLKKNDLEALEROVKELLAONDRDKSKRKQISELEDTATELEQRTVLELE 1531
QY 269 KKSDDPPGNLEPASATNETVSRVLESPAPVEMMNPRLHOPPFQGEIDLNTFTFVNTPT 328
DB 1532 KKQK---NPKILAEKAISEQIAQERDTAREAREKETVLSVSRLEDAFD----- 1581
QY 329 QTSGSQHCPLPKCLERRASPMQNVL-----KKVHKVSKP-----ESQIS-LGGQ 372
DB 1582 -----KIEDLEENKRTLQNELDLDLANTQGTADKNVHELEKAKRALQSQLAELKAQ 1631
QY 373 RCVGELDEEL 382
DB 1632 N-BELEDDL 1639
```

```
RESULT 18
Q9QYT2 PRELIMINARY: PRT: 1447 AA.
ID Q9QYT2
AC Q9QYT2:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE MEA2/GOLGA3 PROTEIN.
GN MEA2/GOLGA3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL6;
RA Matsukuma S.;
RT "Mea2/Golga3 Gene Product is Indispensable for Spermatogenesis in the
RT Mouse.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB029537; BAA66880.2; -;
DR EMBL: AB029521; BAA66890.2; JOINED.
DR EMBL: AB029522; BAA66890.2; JOINED.
DR EMBL: AB029523; BAA66890.2; JOINED.
DR EMBL: AB029524; BAA66890.2; JOINED.
DR EMBL: AB029525; BAA66890.2; JOINED.
DR EMBL: AB029526; BAA66890.2; JOINED.
DR EMBL: AB029527; BAA66890.2; JOINED.
DR EMBL: AB029528; BAA66890.2; JOINED.
DR EMBL: AB029529; BAA66890.2; JOINED.
DR EMBL: AB029530; BAA66890.2; JOINED.
DR EMBL: AB029531; BAA66890.2; JOINED.
DR EMBL: AB029532; BAA66890.2; JOINED.
DR EMBL: AB029533; BAA66890.2; JOINED.
DR EMBL: AB029534; BAA66890.2; JOINED.
DR EMBL: AB029535; BAA66890.2; JOINED.
DR EMBL: AB029536; BAA66890.2; JOINED.
SQ SEQUENCE 1447 AA; 162776 MW; 53087BE7460B3B6F CRC64;
```

```
Query Match 7.9%; Score 189.5; DB 11; Length 1447;
Best Local Similarity 22.7%; Pred. No. 0.023;
Matches 95; Conservative 77; Mismatches 158; Indels 89; Gaps 16;

QY 65 DLAOEENVLDA-EFLKNE-----LDYKQAQLSQKREKR-----DSQAITDRLD 109
DB 712 ELQREADSDREDAIHFLOQEKIVLEVALQSAKSDKEELDRCARLEEDTERTSGILEQLRQ 771
QY 110 TLEENAVESIQNALNKAEMLCSTLKQMKFLERODETQOAREEHR-RODET-----KQ 151
DB 772 DLAVKSNOVEHLQOE-----TATLRKQKQVKEQFVLQGVMEAVTRDRTSKQQLINE 824
QY 152 AREEHRCLKMKTMQIELLLOSQRSEVE---EMIRDMGVGQSAVEQLAVYCVSLKRE 207
DB 825 LKATGKRLRDSMKELRQILKIQGEKKTVVEHSHRLQKMSLVHQQMAELGHLQSVQKE 884
QY 208 YEN-----LKEA-----RKATGELADRLKDKDVSSSKTKTLTNTELDQA 246
DB 885 RDEMEIHLQSLKPFKEQMIATLEANETLKQIELEQDEAKKAITEQKKMKRILSDLSLA 944
QY 247 KLEIRSAQKQSDQADQETSLRKKSDDPPGNLEPASATNETVSRVLESPAPVEMMNPRL 306
DB 945 QKEKTKRKATVEMA---VSIISRQEA---LASKKATDAELNQLRAOSTG---GSSDVL 996
QY 307 HQPPFGDEIDLNTT-----PDVNTPTPTQTSQS-QHCPLPKCLERRASPMQNVLKV 357
DB 997 HEKIRALEVELQNVGQSKILLEKELOEVIYITMSQDELESRRKVLLEBDELOESGFRKXI 1056
QY 358 HKVSKPESQSLGQRCVGVGELDEELAGAPLFIIRNAVIGQKQPNMTTAAESRSSSTDVRI 416
DB 1057 KRLSESNKKLAL-----ELEHERGKLTGIGQSNAAALREHNSILLETVALAKREADIVQL 1108
```

```

RESULT 19
09OYT3 PRELIMINARY; PRT; 1487 AA.
ID Q9OYT3
AC Q9OYT3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MALE ENHANCED ANTIGEN 2/GOLGI AUTOANTIGEN, GOLGIN SUBFAMILY A, 3.
GN MA22/GOLGA3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Matsuoka S.;
RT "Mez2/Golga3 Gene Product is Indispensable for Spermatogenesis in the
RT Mouse."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB029537; BAA86889.2; JOINED.
DR EMBL: AB029521; BAA86889.2; JOINED.
DR EMBL: AB029522; BAA86889.2; JOINED.
DR EMBL: AB029523; BAA86889.2; JOINED.
DR EMBL: AB029524; BAA86889.2; JOINED.
DR EMBL: AB029525; BAA86889.2; JOINED.
DR EMBL: AB029526; BAA86889.2; JOINED.
DR EMBL: AB029527; BAA86889.2; JOINED.
DR EMBL: AB029528; BAA86889.2; JOINED.
DR EMBL: AB029529; BAA86889.2; JOINED.
DR EMBL: AB029530; BAA86889.2; JOINED.
DR EMBL: AB029531; BAA86889.2; JOINED.
DR EMBL: AB029532; BAA86889.2; JOINED.
DR EMBL: AB029533; BAA86889.2; JOINED.
DR EMBL: AB029534; BAA86889.2; JOINED.
DR EMBL: AB029535; BAA86889.2; JOINED.
DR EMBL: AB029536; BAA86889.2; JOINED.
SQ SEQUENCE 1487 AA; 167175 MW; 5BADF716C251FBFF CRC64;

```

```

Query Match 7.9%; Score 189.5; DB 11; Length 1487;
Best Local Similarity 22.7%; Pred. No. 0.023;
Matches 95; Conservative 77; Mismatches 158; Indels 89; Gaps 16;

QY 65 DLAOEENVLDA-EFLKNE-----LDYKAOLQKQREKR-----DSQAIIITLRD 109
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 752 ELQREADSRDAIHFLQNEKIVLEVALQSAKSDKEIDRGARLEDETEFTSGLEQLRQ 811
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 110 TLERNNTVESLQNALKAKEMLCSTLKKQK-KFLQ-----RODET-----KQ 151
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 812 DLAVKSNQVEHLQEE-----TATLRQMKQKVEQFVLQKVMVEAYRRDTSKDLQNE 864
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 152 AREEHAHLKCKMKTMEQIELLQSQSEVE-----EMIRDMGVQSAVEQLAVYCVSLKKE 207
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 865 LKATKRLDSMKRLRBLIKLQGEKTVVEYHSRLQKMSLVQQAELGHLQSQVQKE 924
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 208 YEN-----LKEA-----RKATGELADRLKKDLVSSSKLTLTLELQDA 246
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 925 RDEMEIHLQSLKPEKQMIALTEANETLQKQIEELQGEAKKAITQOKMKRLGSDLTSA 984
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 247 KLELRSAQKDLQSDQETSLRKSSDPPGNLEPASAFTNEVSRVLESPAPVEMMPRL 306
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 985 QKEKTKHKAYENA---VSLSRQLQEA---LASKKATDELNLQRAQSTG--GSSDVL 1036
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 307 HOPFGDEIDLNTT-----PDVNTPTQTQSGS--QHCILPKKLCLEFARSPQNVLKV 357
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1037 HEKIRALEVELQNGQSKILLEKLOEYITMTSQELSESREKVLLEDELQESGFRKI 1096
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 358 HKVAKPESQSLGQRCVGLDELACAFPLFINNAVVGQKQPKRTTAESRSSDVAIRI 416
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1097 KRLRESNKKIAL-----ELEHERGKLTGLGQSNALREHNSILLETALAKREADLVOL 1148
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 20
09H810 PRELIMINARY; PRT; 574 AA.
ID Q9H810
AC Q9H810;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CDNA FLJ13615 FIS, CLONE PLACEL10896, WEAKLY SIMILAR TO NUF1
DE PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Nagatsuna M., Hosojiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NBD0 human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK023677; BAB14636.1;
SQ SEQUENCE 574 AA; 67926 MW; 97EF904DD4EB9EAS CRC64;

```

```

Query Match 7.9%; Score 188.5; DB 4; Length 574;
Best Local Similarity 20.4%; Pred. No. 0.0099;
Matches 80; Conservative 84; Mismatches 147; Indels 81; Gaps 14;

QY 48 QCRIOVQKTIINKL-FFDLAOEENVLDAEFLKNELDYKKAOLQSKDRKRDQAIID 105
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 QAKPENQDKTKVMKLENELEMAQOSAGRDTRFLRNEICQLEKQLEQKDRLEDEMEKELE 139
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 106 ---TLRDTLEARNATVESLQNAL-----NKKEMLCSTLKKQKFLQRODE--TKQA 152
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 140 KKKVNEQLALRNEAEENENSKLRNKKRLKKNQEOICODIIDYQKQIDQSKETILSRG 199
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 153 REEHAHLKCKMKT-----EQIELLQSQSEVEEMIRDMGVQSAVEQ 196
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 200 EDSYRSQSLSKKNTLQYLDLQITLTPANEKITEVQNMKKNLEESVQEA----- 250
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 197 LAVYCVSLKKEYENLKEARKATGELADRLKK-----DLVSSRSK-----LKT 238
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 251 -----EKMTDEYNKMAKIVHQTDNVIDQKKENHYQLOVQELDLKSKNEEDDPIVA 305
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 239 LNTLDAQKLELSAQKDLQSDQETSLRKSSDPPGNLEPASAFTNEVSRVLESPAP 298
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 306 VNAKVEKMKLSSKDEIEIYQOMLHNLREKLN--AQLDADKSNVMAALQCGIOERDSQ 363
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 299 VEMNPNRLHOPFGDEIDLNTT---DVNTPTQTQSGSQCPLKLCLEFARSPQ---NV 353
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 364 IKMLTEQVEO--YTKEMKKNCTIEDLKNELQKNGKA-----STLSQQTTHMKIQSTLDI 415
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 LKKVHKVSKPESQSLGQRCVGLDELACAFPLFINNAVVGQKQPKRTTAESRSSDVAIRI 385
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 416 LKEKTKAEARTAEALAEADAR---EKDKELVEA 444
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 21
09XZE3 PRELIMINARY; PRT; 2138 AA.
ID Q9XZE3
AC Q9XZE3;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MOSIN HEAVY CHAIN.
OS Amoeba proteus (Amoeba).

```

OC Eukaryota; Lobosea; Gymnamoebia; Amoebidae; Amoeba.
OX NCBI_TaxID=5775;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D.
RX MEDLINE=99082369; PubMed=9864850;
RA Oh S.W., Jeon K.W.;
RT "Characterization of myosin heavy chain and its gene in Amoeba
RT proteus".
RL J. Eukaryot. Microbiol. 45:600-605(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=D.
RA Oh S.W., Jeon K.W.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF136711; AAD33718.1;
DR HSP: P08799; IAMD.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR001609; myosin_head.
DR InterPro: IPR004009; Myosin_N.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
SQ SEQUENCE 2138 AA; 244427 MW; DB58B308639FA71 CRC64;

Query Match 7.9%; Score 188.5; DB 5; Length 2138;
Best Local Similarity 22.2%; Pred. No. 0.039;
Matches 102; Conservative 69; Mismatches 170; Indels 119; Gaps 17;

QY 56 KTIINKLFFDLAQBENVLDAFLKNEIDSKAQLSQ----- 92
DB 1247 KRLVAKLESDKKMKKEVEDERDLKNKLDQK-KLSQALDELGNALKEAMKRSREKN 1305
QY 93 -KDREKR-----DSQAIITDPL-----RDTLEERNATVESLONALKAMKLCSTL 135
DB 1306 RKDLENRLRELEDDQAEQAAARSNEKEFRGFEDNLEHQSDVDEVVLSAAKKKL 1365
QY 136 KKQMKLEFQODETKQAREAHRLCKMKMTME---QIELLQSQSEVEEMIRMGVGO 191
DB 1366 ESELEAKRSLDNEAGEKVAEE---KMKVLDTELHEQLALSNANKNTGLVRVKKVQ 1422
QY 192 SAVQGLAVVYCVSLKKEYENLKEARKATGELADRLKKDLYSSRSKIKTINTELDQAKELR 251
DB 1423 DEVEDLNEQYEMASKELSKLDKGNKKTAELEKELRRHVQESQS-----SLDAGELKLR 1475
QY 252 SAQKDLQSAQDEITSLRKKS-----DDPPGNLEPASATNETVSRILVFESPA 297
DB 1476 HTQDELDELHNLQJEDLEKSSLSRSKKQLOLYDDLEDTEHEELAATKAEKRLVKDLEA 1535
QY 298 PV-EMANRHLQPPFGD-----EIDL-----NTTF-----DV 323
DB 1536 DIALEQETRVSEFPLMQAEKALSLLEVELVLDLKKADQSOAFKAVENERSALREYBDL 1595
QY 324 NTPPTQTS---GSGHCLPKKL-----CLERARSPMOWNLKVVHVSREPQSLISGG 371
DB 1596 QAOQDETSKNIANDRAKRIKNTLDQELSKLEKASNAQKLEKRLKKAERDLAAKAAAS 1655
QY 372 QRCVGEI-DEEL--AGAFPLFIRNAVIGQKOPNRTTAESR 408
DB 1656 ARAGGCVSDELRRAQAEALALRDA-DREKSNKITAEKR 1694

RESULT 22
ID 028714 PRELIMINARY; PRT; 1156 AA.
AC 028714;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CHROMOSOME SEGREGATION PROTEIN (SMC1).
GN AFI1558.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrtides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sultion G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glöck A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Arlrich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus".
RL Nature 390:364-370(1997).
DR EMBL: AE000995; AAB89690.1;
DR TIGR: AF1558;
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR003405; SMC_C.
DR InterPro: IPR003395; SMC_N.
DR InterPro: IPR02017; Spectrin.
DR Pfam: PF02483; SMC_C; 1.
DR Pfam: PF02463; SMC_N; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; UNKNOWN_1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 1156 AA; 134244 MW; 18CB16F7CD5AD1ED CRC64;

Query Match 7.8%; Score 186.5; DB 17; Length 1156;
Best Local Similarity 16.7%; Pred. No. 0.026;
Matches 78; Conservative 111; Mismatches 152; Indels 125; Gaps 13;

QY 55 KTIINKLFFDLAQBENVL-----DAEFLKNE----- 82
DB 671 KERMISDKIYELQREKEGLFAELNRAESLRQYQDEVORLGTWISLNRISLDEKIRT 730
QY 83 ----LDSKADLSQDKRK-----RDSQAIITDPLRDTLEERNATVESLONALKAMK 131
DB 731 ESGRIEELREKISQSRKENYISLKDYNKSLAMEBAIGLEAEIEIERMLRGSE-- 788
QY 132 CSTLKKQMKFLEQODETKQAREAHRLCKMKMTM---EQIELLQSQ-----RSEV 180
DB 789 ---VPKIVIEELDKIKEEQRNREILISTEKKIESIEFRREDDLESSMOEKQOYDEIDRI 845
QY 181 EEMIRDMGVQSAVEQLAVVYCVSLKKEYENLKEARKATGELADRLKKDLYSSRSKLTN 240
DB 846 DEIRRTIEEGKARAVEINSELEIRKEBERELGELGKRGREDILQLRNAEERKRIE 905
QY 241 TELDQAKLELSAQKDLQSAQDEITSLRKKSDDPPGNLEPASATNETVSRILVFESPAVE 300
DB 906 AEIRLERIRIKLOREKRIEIAESEAIGEV--EVPENLPLEKVEKYLDELV----- 955
QY 301 MMNRLHQPFGDELDINTFTDVMTPTQ-----TSGSGHCLPKKLCLERARSPMOWNLK 355
DB 956 -----VELSTFGDVNLKALQIEEYKARRDELVEKKNVLEKERA---DILD 998
QY 356 KVRVSRPESQSLISGQRCVGEIDELDELAGAPLFIIRNAVIGQKOPNRTTAESRSTQVR 415
DB 999 RIEKYEKMKREI-----FEVE-----TAINRFAELIR 1027
QY 416 IGFGLGGRKTFQPRDTTIRPVVSKAKSKOKVRIRKITYSSASQ 461
DB 1028 ---ELANQEGELYLDSDDPFNSGLYIKVYKPNKPVQKLSKSGEK 1070

```

RESULT 23
ID 026079 PRELIMINARY: PRT: 1941 AA.
AC 026079:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MYOSIN HEAVY CHAIN.
OS Placopecten magellanicus (Sea scallop).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoida;
OC Pectinoidea; Pectinidae; Placopecten.
OX NCBI_TaxID=6577;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKLETAL MUSCLE;
RA Perreault-Micale C., Kalabokis V., Nyitrai L., Szent-Gyorgyi A.G.;
RT "Sequence variations in the surface loop near the nucleotide binding
RT site modulate the ATP turnover rates of molluscan myosins."
RL J. Muscle Res. Cell Motil. 0:0-0(1996).
DR EMBL: U59294; AAB03660.1; -.
DR HSSP: P24733; IMDC.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR001609; myosin_head.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR Pfam: PF00612; IQ_2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS00096; IQ; 1.
DR PROSITE: PS00096; IQ; 1.
SQ SEQUENCE 1941 AA; 223243 MW; E2AD1637FFDB8127 CRC64;

Query Match 7.7%; Score 184; DB 5; Length 1941;
Best Local Similarity 22.4%; Pred. No. 0.061;
Matches 97; Conservative 71; Mismatches 131; Indels 134; Gaps 19;

QY 55 KRIIINKLPFD--LAQEEENVLADEFKLNELDSVK-----AQLSQ 92
DB 906 EKLIQKADPESSQIKLEERLDEEDAAADLEGIKKMEGDANLNKKDIGELEHSLQSE 965
QY 93 KDRKKRDSQAIIIDL-----RDITLERN-ATVESLQ-----NALNKA 128
DB 966 QERAHKDNQ--ISTLOGESQDEHIGLKKKEKKALEFNAKKTSDSLQAEEDKCNHLNK- 1022
QY 129 EMLCSTLKQMKFLEORODETKQAREEHNRLKCKMK-----TWEQIELLLQSQRSEVEEM 183
DB 1023 --LKAKLEQALDELEDLREKKKVGDEKAKSKVEDLKSQENVEDLERVKR-EELEN 1079
QY 184 IR-----DMGVGSAVQQLAVYCVSLKKEYENLKEAKKATGELADRLKDLVSS 232
DB 1080 VRKKEAEITTLNSKLEDEQNLVSOLQRIKELQARIEELEEELEAERNASVKEKQRAEL 1139
QY 233 RSKRLKTLNLELDA-----KLELRSAQKDLQSA-----DOETISLRKKSDDP 274
DB 1140 NRELEELGEHLDEAGATSAQIQLNKKREAEELKIRDLDEEASIQHNAQISAIRKKKQD- 1198
QY 275 PGNLEPASATNETVSRLVFESSPAVEEMNRLHOPPGDEIDLNTTFDVNTPTQTSGSQ 334
DB 1199 -----AANEMADGV-----DOLQVKSLEK-----DKDKLREMD-----DLESQMT 1236
QY 335 HCLPKKLCLEARRSPMGNVLKVKHVKSPESQSLSGQRCVGEDELDELACAFPLFTINAV 394
DB 1237 HDKNNKGSCKSVKQKQFES-----QVSDLNARLE-DSQRSINEL----- 1273
QY 395 LGOKOPNRRTAES 407

```

```

DB 1274 --QSKSRLOAEN 1284
RESULT 24
ID 091BD4 PRELIMINARY: PRT: 1937 AA.
AC 091BD4:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MYOSIN HEAVY CHAIN.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=HEART;
RX MEDLINE=20149854; PubMed=10684978;
RA Machida S., Noda S., Fututani Y., Takao A., Momma K., Matsuoka R.;
RT "Complete sequence and characterization of chick ventricular myosin
RT heavy chain in the developing atria."
RL Biochim. Biophys. Acta 1490:333-341(2000).
DR EMBL: AB032197; BAA92710.1; -.
DR HSSP: P08799; 1MND.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR001609; myosin_head.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR000533; Tropomyosin.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRINTS: PR00194; TROPOMYOSIN.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
SQ SEQUENCE 1937 AA; 223473 MW; 95A94F69CDB36781 CRC64;

Query Match 7.7%; Score 183.5; DB 13; Length 1937;
Best Local Similarity 20.4%; Pred. No. 0.065;
Matches 78; Conservative 81; Mismatches 165; Indels 59; Gaps 11;

QY 59 INKLFEDLAQEEENVL--DAEFKLNELDSYKAOLSKDRKKRDSQAIIIDLRLDLEERNA 116
DB 1267 MTRLMNDLTQKTRKLOSENSEFVR-QLEKESLSLSQSRKSTFTQOIELRLRQLEERTK 1325
QY 117 TVESLONALNKAMLCSTLKQMKFLEORODETKQAREEHNRLKCKMKKTWEQIELLLQSQ 176
DB 1326 SKNALNALQAAAHDDLDLREQYEEEOEAKAELOALSKGNAEVAQWRTYETDAIORT 1385
QY 177 RSEVEEMIRDMGV-----GQSAVEQQLAVYCVSLKKEYENLK-----EARKATGELADR 224
DB 1386 --ELEDKAKKLLAARLOEAEEAITEANAKCSLSKTYNRLQNELEDMMITLKNNSAAAL 1443
QY 225 LKRDVSSRSKLTN---TELDOAKLELRSQAOKDLQSAQDOETISLRKKSDDPPGNLEPA 281
DB 1444 DKK---QRFEDKTIINDWKQKYESQAELEASQKEARSLSLELFKLNVAEETLDHETL 1499
QY 282 SATNETVSRLVFESSPAVEEMNRLHOPPGDEIDLNTTFDVNTPTQTSGSQHCLPKKL 341
DB 1500 KRENKNLQOEIISDLTNOISGNKNLH-----ETL-----KVKQVDEKSEVOL 1543
QY 342 CLRARRSPMGNVLKVKHVKSPESQSLSGQRCVGEDELDELACAFPLFTINAVIGOKOPN 401
DB 1544 ALDEAGALDEHESKTLRFOLELSQLKADPERKLADEDEMEI-----RRNQ 1592
QY 402 RT-----TASRSSTDVARI 416

```

Db 1593 RTIDSLDSTDEARSREAIRL.1615

RESULT 25

Q9NJ23 PRELIMINARY; PRT; 1219 AA.

AC Q9NJ23;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MYOSIN HEAVY CHAIN STRIATED MUSCLE SPECIFIC ISOFORM (FRAGMENT).
GN MHC.
OS Aequipecten irradians (Bay scallop).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoida;
OC Pectinoidae; Pectinidae; Argopecten.
OX NCBI_Taxid=31199;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20090924; PubMed=10623517;
RA Yamada A., Yoshio M., Ojwa K., Nyitray L.;
RT "catlin, a novel protein in molluscan catch muscles, is produced by
RT alternative splicing from the myosin heavy chain gene.";
RL J. Mol. Biol. 295:169-178(2000).

DR EMBL: AF183909; AAF62391.1; -;
DR HSSP: P24733; IWDG.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR001609; myosin_head.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR000533; Tropomyosin.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00194; TROPOMYOSIN.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR PROSITE: PSS0096; IQ; 1.
FT NON_TER 1 1
SQ SEQUENCE 1219 AA; 141202 MW; BAC6F574DA4AFBBB CRC64;

Query Match 7.6%; Score 183; DB 5; Length 1219;

Best Local Similarity 23.0%; Pred. No. 0.042; Matches 102; Conservative 67; Mismatches 128; Indels 146; Gaps 20;

QY 55 KTIINKLFPD-----LAOEENYLDSEFLKNEIDSVKADL-----SQ 92
DB 184 EKLIMQKADDESQIKLEBERLDEEDAAADLEGITKKMEADNANLKKDIDGENTLOKAE 243
QY 93 KDRKRDQAIIIDL-----RDTLEERN-ATVESIQ-----NALNKA 128
DB 244 QDKAHKNO--ISFLQGEISQDEHIGKLNKKKALEANNTKTSIDLAEDKCNHLNK- 300
QY 129 EMLCSTLKKQKFLQKODEFKQAREAHNLKCKMK-----TMDQIELLOSQSEVEEM 183
DB 301 --LKALEQALDELEDEMLERKKVRGDEVKAKRVEODLKSTQENVEDLEKVR-ELEEN 357
QY 184 IR-----DMVGQSAVEQLAVYCVSLKKEYENLKARATGELARLKKDLVSS 232
DB 358 VRKKEAFISSNKLDEQDNLVSQLRKIKELQRIELEELEAEARNARAKVKQRAEL 417
QY 233 RSKILNTLTDQA-----KLELRSAQKDLQSA-----DOEITSLRKKSDP 274
DB 418 NRELEELGERLDEAGATSAQIELNKKRREALLKIRDLERASQIHQNSALKKKHOD- 476
QY 275 PGNLEPASATNETVSRVLFESPAPVEMNPRLOPPFGDEIDLTTPTDVTPTQTSQSO 334
DB 477 -----AANEMADQY-----DQLQKVKSKLEK-----DKKDLKREMD-----DLESQMT 514
QY 335 HCLPKKICLERARSPMONVLKKVHKVSKPRESOLS-----LGGRCVGEIDELAGAPPL 388
DB 515 HNMKNKGCSEK-----VMKQF-----BSQMSDLNARLEDSQRSINEL----- 551

QY 389 FIRNAVLGOKOPNRTTAESRSST 411

Db 552 -----QSOKSRLOAENSIDL 566

RESULT 26

Q9NCL3 PRELIMINARY; PRT; 1871 AA.

AC Q9NCL3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MUD PROTEIN.
GN MUD.
OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20345122; PubMed=10884435;
RA Guan Z., Prado A., Melzig J., Heisenberg M., Nash H.A., Raabe T.;
RT "Mushroom body defect, a gene involved in the control of neuroblast
RT proliferation in Drosophila, encodes a coiled-coil protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:8122-8127(2000).

RN [2]
RP SEQUENCE FROM N.A.
RA Prado A., Raabe T.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF209068; AAF89163.2; -;
DR FlyBase: FBgn0002873; mud.
SQ SEQUENCE 1871 AA; 216126 MW; AE3BE89F937F76D CRC64;

Query Match 7.6%; Score 181.5; DB 5; Length 1871;

Best Local Similarity 22.0%; Pred. No. 0.08; Matches 79; Conservative 69; Mismatches 144; Indels 67; Gaps 11;

QY 56 KTIIT---NKLFDLAOEENYLD--AEFLKNEIDSVKQKQSDK-----REK 97
DB 1243 KSVIEAQTQKSDLDQREKESAOQLVDNLKVELDERKRELQAVNSAFEAQTKSDLOREK 1302
QY 98 RDSQAIIIDLTLDEERNATVESLONALNKAKEMCSTLKKOMKFLQRODETKOAREAH 157
DB 1303 ESAQOVLVDNLKVELDKERKELAQVKSVEAQTQKSDLDQREKESAOQLVDNLK- 1356
QY 158 RLCCKMKTQEIILLQSORSEVEEMIRDMGVGQSAVEQLAVYCVSLKKEYENLKARKA 217
DB 1357 ELDERKRELAKVKSVIEAQTQKSDLDQREKESAOQLVDNLKVELDKERKELAQVNSAFEA 1416
QY 218 TGEIADLKLKDVSSRSKLTNLTELQAKIELRSQKDLQSAQDETSLRKKSDPPGN 277
DB 1417 QTKLSDDLQKESAOQLVDNLKVELDKER-----KELAQVSAIGAQTQKSD- 1465
QY 278 LEPASATNETVSRVLFESPAPVEMNPRLOPPFGDEIDLTTPTDVTPTQTSQSOHCL 337
DB 1466 ---LEQKESYQOQLVDNLKVELDKERELAK-----VNSAFEAQTK-----L 1504
QY 338 PKKLCLERARSPMONVLKKVHKV-SKRESQSLG-----GRCVGEIDELAGAF 386
DB 1505 SDDLKQKEDAQREVFLVKRELVEKKEFEVKLATLTDIIETLEMRCT-QMEEERATAY 1562

RESULT 27

Q9EON8 PRELIMINARY; PRT; 324 AA.

AC Q9EON8;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE MITOSIN (FRAGMENT).
OS Mus musculus (Mouse).

DR EMBL: D50474; BAA09067.1; -
 DR HSSP: P13538; 2MYS.
 DR InterPro: IPR000048; IO.
 DR InterPro: IPR001609; myosin_head.
 DR InterPro: IPR004009; myosin_N.
 DR InterPro: IPR002928; myosin_tail.
 DR Pfam: PF00612; IO; 1.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF02736; myosin_N; 1.
 DR Pfam: PF01576; myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PSS0096; IO; 1.
 KW myosin.
 SQ SEQUENCE 1933 AA; 221092 MW; CDF0CBAA475530F5 CRC64;

Query Match 7.6%; Score 181; DB 13; Length 1933;
 Best Local Similarity 19.9%; Pred. No. 0.088;

Matches 99; Conservative 91; Mismatches 170; Indels 138; Gaps 18;

QY 31 LQCLIQMFETAPSRTPCPCRIQVGKTIINKLFEDLAQEEENVLD----AEFLKNELDS 85
 DB 1142 LEETISERLEEGAGTAQAIEMKNKREAEFOKMRDL---EESTLQHEATAALRRKKQADS 1198
 QY 86 VKAQLSOK-----DREKDSQAIITPLRDLTEERNAVESLONALKNKEMLCST 134
 DB 1199 V-AELGEIDNLQRIKOKLEKESEYKMEIDL-----SSNMEAVAKAGNLEKMCRT 1250
 QY 135 LKKMKFLEORODETKQAREAHRLCKMKT-----MEOIELL----- 172
 DB 1251 LEDDLSELKAKSDENSRLNOMNARALQTENGEFSQLEKEALVSQLTRGQAFTQ 1310
 QY 173 LOSRSEVEEMIRMGVQSAVEQLAVYCVSLKREYENLKAR-----KATGELAD- 223
 DB 1311 IEDLKRHYEEEVAKAMLAHVAOSARHDCDLLRQEEEOQAKAELQSGMSKANSEVAKOW 1370
 QY 224 -----RLKKDIVSSRSKLTINTE---LQAKLELSAOKDL--- 257
 DB 1371 RAKETDAIOTTELEESKKKLAQLQDAEESIEAVNSKCSLSETKQRLQSEVEDLMD 1430
 QY 258 -----OSADEITSLRKSSDDPKNLEPASATNFTVSLVESPAPVE-- 300
 DB 1431 GERANALANLDKKQRNDDKVLADWKQYEEQAELQAQKARSLSLELFKMNSTYEEA 1490
 QY 301 -----MANPRLHQPPFGDEIDLNTTFDVNTPPTQTSQSQCPLP-----K 340
 DB 1491 LDHLETLKREKNLQO-----EIS-----DLSEQLGETGKSIHEIEKAKKTVESEKAEIQ 1540
 QY 341 LCLERARSPMONVLKRYHKVSKPESQSLSGQRCVGEDELDELAGAFLFTINA--VLGQK 398
 DB 1541 TALEEAGTLEHESKTLRLQLELNOVKSETIDRLAEKDEEM---EQIKNSQRLVDSM 1596
 QY 399 QPNRTAESRSSIDYVRI 416
 DB 1597 Q-STLSEVRSRNDALRV 1613

RESULT 30
 Q9UT73 PRELIMINARY; PRT; 1940 AA.
 AC Q9UT73;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE MYOSIN HEAVY CHAIN.
 OS Pecten maximus (King scallop) (P.ligum's clam).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Periomorpha; Pectinoidea;
 OC Pectinoidea; Pectinidae; Pecten.
 OX NCBI_TaxID=6579;
 RN [1]

RP SEQUENCE FROM N. A.
 RC TISSUE=ADDUCTOR MUSCLE;
 RA James D.P., Patel H., Chantler P.D.;
 RT "Primary structure of myosin from the striated adductor muscle of the
 RT Atlantic scallop, Pecten maximus."
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF134172; AAC52842.1; -
 DR HSSP: P24733; IMDC.
 DR InterPro: IPR000048; IO.
 DR InterPro: IPR001609; myosin_head.
 DR InterPro: IPR004009; myosin_N.
 DR InterPro: IPR002928; myosin_tail.
 DR InterPro: IPR002017; Spectrin.
 DR Pfam: PF00612; IO; 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF02736; myosin_N; 1.
 DR Pfam: PF01576; myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IO; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PSS0096; IO; 1.
 SQ SEQUENCE 1940 AA; 223141 MW; A3D09DECBEB16F90 CRC64;

Query Match 7.6%; Score 181; DB 5; Length 1940;
 Best Local Similarity 23.0%; Pred. No. 0.088;

Matches 101; Conservative 67; Mismatches 125; Indels 146; Gaps 20;

QY 55 KTIYINKLFED--LAQEEENVLAQEFKNELDVSK-----AQLSQ 92
 DB 906 EKLIMQKADFESQIKLELERLDEBDAADLEGIKKMETNSLNKKDIDGLENTIQAE 965
 QY 93 KOREKDSQAIITPL-----RPTLEERN-ATVESIQ-----NALNKA 128
 DB 966 QDKAHKDNQ--ISTLOGEMAQODEHIGIKLNKKEKALAEANKKTSLSQAEEDKCNHLNK- 1022
 QY 129 EMLCSTLKKMKFLEORODETKQAREAHRLCKMKT-----TMOIELLQSORSEVEEM 183
 DB 1023 --LKALEQLADELDELNLKERKKVGVKAKKRVBODLKTQENVEDLEVKR-ELLEN 1079
 QY 184 IR-----DMGVQSAVEQLAVYCVSLKREYENLKARATGELADRLKDIVSS 232
 DB 1080 VRKREAEISTLSKLEDBQNLVSOQRKVKELQRIELEBELEERNARAKVEKQRAEL 1139
 QY 233 RSKLTNLTEDQA-----KLELSAQKDLQSA---DOETLSLKKSDDP 274
 DB 1140 NRELLEELGERLDEAGATSAQIELNKKREAEILKIRDLLEASLQHEAOISALRRKHQD- 1198
 QY 275 PGNLEPASATNETVSRILVFESPAPVEMMNPRLHQPPFGDEIDLNTTFDVNTPPTQTSQSQ 334
 DB 1199 -----ANEMADQV-----DQIQVKSYLEK---DKQDIREAD-----DLESQMT 1236
 QY 335 HCLPKLCLEERARSPMONVLKRYHKVSKPESQSL-----LGGRCVGEDELDELAGAFLP 388
 DB 1237 HNMKNKGCSEK-----VMKF-----ESQMSDLNARLDSQRSINEL----- 1273
 QY 389 FIRNAVLGQKQPNRTAES 407
 DB 1274 -----QSQKSRLOAEN 1284

RESULT 31
 Q02015 PRELIMINARY; PRT; 2007 AA.
 AC Q02015;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE NONMUSCLE MYOSIN HEAVY CHAIN.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Callus.
 RX NCBI_TaxId=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9238144; PubMed=1355479;
 RA Takahashi M., Kawamoto S., Adelstein R.S.;
 RT "Evidence for inserted sequences in the head region of nonmuscle
 RT myosin specific to the nervous system. Cloning of the cDNA encoding
 RT the myosin heavy chain-B isoform of vertebrate nonmuscle myosin.";
 RL J. Biol. Chem. 267:17864-17871(1992).
 DR EMBL: M93676; AAA48987.1; -
 DR EMBL: M93676; AAA48986.1; -
 DR EMBL: M93676; AAA48985.1; -
 DR EMBL: M93676; AAA48988.1; -
 DR HSSP: P10587; 1BR2.
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR001609; myosin_head.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR002017; Spectrin.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF00663; myosin_head; 2.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 DR PROSITE: PS50096; IQ; 1.
 SQ SEQUENCE 2007 AA; 232592 MW; 36A591368CD264BA CRC64;

Query Match 7.68; Score 181; DB 13; Length 2007;
 Best Local Similarity 21.58; Pred. No. 0.091;
 Matches 96; Conservative 87; Mismatches 167; Indels 96; Gaps 16;
 QY 68 QEEENVLDIAEFLKNEIDSVKALQSKDKREKDSQAIIIDTLTLEERNATVESIQNALNK 129
 DB 1536 QNKRLADMEDLMSSKDDVGNVHLEKSKR-----TLEQVEEMTQLELEDELQA 1588
 QY 128 AEMLCSTLKQMKLEORODETQAREBAHRLCKM--KTMEQIELLIQSQRSE----- 179.
 DB 1589 TEDAKLRLEVMMQAKKQFEDLDQARDEONEEKRMLVKQVRELAELERKORALAVA 1648
 QY 180 -----VEEIRDMVGO-----SAVEOLAVYCVSLKREYMKERKATGELADRLK 226
 DB 1649 AKKMEHMDKDL-EGQLEANKARDEAIKQRLQAQMKDYORELEEARASRDEIFPAQSK 1707
 QY 227 KDLYSSRSKLTNTLTDQAKLELSAQKDLQSAQDOETISLRKSDPFGNLEPASATNE 286
 DB 1708 E-----SEKTLKGLAEITLQEQEFAASERARHHAQDEDEL--ADELIANSASGKSALLD 1760
 QY 287 TVSLV-----VESPAVPEMMANRHLHQPFGDEIDINTTFDVNTPTPTQTSQSQC 336
 DB 1761 EKRRLEARIQLEBELEEDSNNMELNERFRK-----TTLQVDTLNLSELAGERSA 1810
 QY 337 LPKLTCLERASPMQNVLLKYVH-KVSKPESQSLSGQKCVGELDELAGAPPLIRNAVVL 395
 DB 1811 AQKS-----ENARQOLERONKELAKLQLEGSVKSKFKATITLLEKTIQL----- 1857
 QY 396 GQKQPNRTTASRSSTDVVRIGFDLGGRTK-----FIQPRD-----TTIIRPVVKSKA 445
 DB 1858 -EEGLEDAKRAAANKLVK-----RTEKKIKVEYMWQVEDERHNDQYKKEQMEKANA 1908
 QY 446 KSKQKVR-----IKYVSSAQPKL 464
 DB 1909 RMKQLKROLEAEAEATRANASRRKL 1934

RESULT 32
 ID Q96720 PRELIMINARY; PRT; 692 AA.
 AC Q96720;

DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PARAMOSTIN-LIKE ALLERGEN (FRAGMENT).
 OS Dermatophagoides farinae (House-dust mite).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
 OC Acariformes; Sarcophormes; Astigmata; Analgoidea; Pyroglyphidae;
 OC Dermatophagoides.
 OC NCBI_TaxId=6954;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98389068; PubMed=9723675;
 RA Tsai L.C., Chao P.L., Shen H.D., Tang R.B., Chang T.C., Chang Z.N.,
 RA Hung M.W., Lee B.L., Chua K.Y.;
 RT "Isolation and characterization of a novel 98-kd Dermatophagoides
 RT farinae mite allergen.";
 RL J. Allergy Clin. Immunol. 102:295-303(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20062524; PubMed=10594536;
 RA Tsai L., Sun Y., Chao P., Ng H., Hung M., Hsieh K., Liaw S., Chua K.;
 RT "Sequence analysis and expression of a cDNA clone encoding a 98-kda
 RT allergen in Dermatophagoides farinae.";
 RL Clin. Exp. Allergy 29:1606-1613(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Tsai L.-C., Chao P.-L., Ng H.-P., Chua K.-Y.;
 RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF352244; AAK39511.1; -
 FT NON_TER 692
 FT 692
 SQ SEQUENCE 692 AA; 81372 MW; 61FC6380C9D7C9E2 CRC64;

Query Match 7.58; Score 180; DB 5; Length 692;
 Best Local Similarity 21.98; Pred. No. 0.034;
 Matches 87; Conservative 64; Mismatches 160; Indels 86; Gaps 13;
 QY 59 INKLEFDIAEENVL--DAEFLKNEIDSVKALQSKDKREKDSQAIIIDTLRDTLEERNA 116
 DB 104 INRVIELTSHKQRLSQENTELIK-EVHEVQLQDNNANHLKTQIAQOLEDRHLEEEER 162
 QY 117 TVESIQNALNKAEIMCSTLKQMKLEORODETQAREBAHRLCK-----MKTMEQIEL 171
 DB 163 KRASLENAHTLEVELESLKQVL-----DESEARLELRQTLKANGDAAMSKSKEYA 215
 QY 172 LLQSORSEVEEMIDMGV-----QSAVQALAVYCVSLKREYENK-----EARRATG 219
 DB 216 ELQAHAEVEEELRRKMAQKISEYEEOLEALLNKCSLEKQSRLOSEVEVLIIMLEKATR 275
 QY 220 ELAD-----RLKDDVSSRSKLTNTLTDQAKLELSAQKDLQSAQDOETISLRKSSD 273
 DB 276 HAQOLEKRVADLEKINIDLKRLKEEYVLMQDAQKELRYKIAELQKLOHEEKLRDQDQ 335
 QY 274 PPGNLEPASATNETVSRVSPAPVEMMNRLHQPFGDEIDINTTFDVNTPTPTQSGS 333
 DB 336 -----LARENKLTLDLAQKSLQNDARRIH-----QELIEIKR----- 370
 QY 371 -----LENERDELSAAYKAEETLKRQEBAN--QRLIAELAVRNDYERLAKOD 418
 DB 380 ELAGAPFLTRNAVIGQKQPNRTTASRSSTDVVR 416
 QY 419 EIEELRKQY--QTEIEQLNMRLAEAEKLTETIARL 453

RESULT 33
 ID Q21440 PRELIMINARY; PRT; 1992 AA.
 AC Q21440;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)

Db 1410 ALETAGSKCLHDTKNQILOEELLNKTQVKCEKLNKK-----KLBOEVI 1458
QY 384 GAFFLIRNAV-LGOKOPNTTAESSRSTPDVVRIGFDGLGRTKFIOP-----R 431
Db 1459 NLRSHIERNNVELGQYKQYQOEIEERARQEI-----EKLKEVNLFLQAOAASQENLEQFR 1514
QY 432 DTTIIRPVYKSAKSKQVRIKTSS 458
Db 1515 ENNF-----ASMSQOMELRIKDL 1534

RESULT 35
Q9H1Q1 PRELIMINARY; PRT: 1710 AA.
ID 09H1Q1
AC 09H1Q1
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE BAI45E8.1 (KIMA1074).
GN BAI45E8.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eultheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird C;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162272; CAC19649.1; -
DR InterPro: IPR002110; ANK.
DR Pfam: PF00023; ank; 4.
DR SMART: SM00248; ANK; 4.
DR PROSITE: PS0297; ANK_REPEAT; 4.
DR PROSITE: PS0297; ANK_REPEAT; 1.
KW ANK repeat; Repeat.
KW ANK repeat; Repeat.
SQ SEQUENCE 1710 AA; 196410 MW; 01CBP9BADB894872 CRC64;

Query Match 7.5%; Score 179.5; DB 4; Length 1710;
Best Local Similarity 19.2%; Pred. No. 0.093;
Matches 109; Conservative 111; Mismatches 182; Indels 165; Gaps 22;

QY 16 HSRDVAIHGHTFHLOCLIQWETAPSRTPQC-RIOVGKTIINKLFFDLAQBEEEN-- 72
Db 1010 HSRDLAAIH--DRDQSTSRKLELAFGRARDCSRQ-----DKMNDVSLKKNNE 1060
QY 73 -----VDAEFLKNELDVYKALOSKDKREKDSQA 102
Db 1061 ILSQQLFTESKLSLEIEFHHTDALREKTLGLERYQKDLSTOCOMKEQKYQEQV 1120
QY 103 IIDTL--RDTLEERNATVES-----LQNLNKA-----EMLGS 133
Db 1121 KVKYTGKQSVBERLSQLOSEMMLRQOLDHNRKADNKEKTYINIQDFHAIYQKLA 1180
QY 134 TLKQKFLERODETKQAREEAHRLCKMKMTW-----QTELLLOSQSEVEEMIRDMGV 189
Db 1181 ESKQSLLEER---NKELISECNHLKERQYQYENENKAKEBEVYVROQLADLTLLKQSM 1237
QY 190 GQSAVEQLAVYCVSLKKEYNLKARKATGELADRL-----KKDL 229
Db 1238 SEASLEVTSTYRINLEDETODLK---KLGQIRNQLQLEADRHTEAVRCAEKKMDHKOKL 1294
QY 230 VSASRSLK-LTLNTELDQAKLELSAQKLOLSAD-----OEITSLRKSSDDPGNLEPAS 282
Db 1295 EKDNALKTIVKQMD---KIELOKLNLANLSDEKQKLMELKOSLLECNDQEM 1350
QY 283 ATNETVSRLVFESPAPYEMNPRLOPPFGD-----EID-----LNTTFDVT 325
Db 1351 KKNVLEIREITGFKNLKMKRKLNENEENEFSPHGLKTSQFEMDIQIKLKHKIDLT 1410
QY 326 PPIQTSQO--HCLPKKLCERARSPQONLVKVKHYKSPESQSLSGQRCVGLDELA 383
Db 1411 ALETAGSKCLHDTKNQILOEELLNKTQVKCEKLNKK-----KLEQVI 1459

QY 384 GAFFLIRNAV-LGOKOPNTTAESSRSTPDVVRIGFDGLGRTKFIOP-----R 431
Db 1460 NLRSHIERNNVELGQYKQYQOEIEERARQEI-----EKLKEVNLFLQAOAASQENLEQFR 1515
QY 432 DTTIIRPVYKSAKSKQVRIKTSS 458
Db 1516 ENNF-----ASMSQOMELRIKDL 1535

RESULT 36
Q90338 PRELIMINARY; PRT: 1092 AA.
ID Q90338
AC Q90338;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-JUN-1998 (Tremblrel. 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MYOSIN HEAVY CHAIN (FRAGMENT).
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Cyprinus.
OX NCBI_TaxId=7962;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FAST MUSCLE;
RX MEDLINE=97176447; PubMed=9023993;
RA Imai J., Hirayama Y., Kikuchi K., Kakinuma M., Watabe S.;
RT "CDNA cloning of myosin heavy chain isoforms from carp fast skeletal
RT muscle and their gene expression associated with temperature
RT acclimation.";
RL J. Exp. Biol. 200; 27-34(1997).
DR EMBL; D50475; BA09068.1; -
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR000533; Tropomyosin.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00194; TROPOMYOSIN.
FT NON TER 1
SQ SEQUENCE 1092 AA; 125885 MW; B6BABA3963BEEBA CRC64;

Query Match 7.5%; Score 179; DB 13; Length 1092;
Best Local Similarity 17.7%; Pred. No. 0.062;
Matches 89; Conservative 94; Mismatches 172; Indels 148; Gaps 13;

QY 31 LQCLIQWETAPSRTPQCRIQYGKTIINKLFFDLAQBEEENVLDAEF----- 78
Db 302 LLEISERLEEGAGATAQIIMNRREADFQKMRDL---BESTLQHEATTAALRKQADT 358
QY 79 ---LKNELDVYKALOSKDKREKDSQAIIDTLRDTLEERNATVESLQNALNKAEMLCSTL 135
Db 359 VAELEGQIDMLQYKQKLEKESEYKMEIDL-----TSNMEYAKAKGNLEKMRCL 411
QY 136 KQKFLERODETKQAREEAHRLCKMKT-----MEQIEL-----L 173
Db 412 EDQSEIKAKSDNSRQNDNMNQARLQTEENGFSQLEBEKALVQLTRGQAFTQI 471
QY 174 QSORSEVEEMIRDMGVQSAVEQLAVYCVSLKKEYNLKARK-----KATGELA--- 222
Db 472 EDLQHVHEEVKAKMALAHVQSA RHDCDLRQYEEQDEAKTELQSGMSKANSEVAQWR 531
QY 223 -----DRLKTL-----VSSR----- 233
Db 532 AKYETDAIOTTELEBEKAKKLAORLQDAESIYAVSSKASLEKTORLOGEVEDLMTIG 591
QY 234 -----SKLKTNLTELDQAKLELSAQKLOLSADQDITSLRKKSDPP 275
Db 592 ERANANLKKQORNDKVLAEWKQANAKYEESSQALEBAQKEARSLSTELFKMNSYEAL 651
QY 276 GNLSPASATNETVSRLVFESPAPYEMNPRLOPPFGDELIDLNTTFDVTNPPQTSSQO 335
Db 652 DHELTREKRNKNLQOEISLSEQLGETGKSIHELEKAKKIVSEKAEIQT----- 701

QY 336 CLPRLKLERARSPMGNVLKVKHKVSKPESQSLSGQRCVGEDELDELACGAPLFLIRNA-- 393
 DB 702 -----ALEEAESELEHEESKILRVQLELNGVKSEIKRLAKNEEL-----EQIKRNSQR 751
 QY 394 VLGGKOPNRTTAESRSSTDVRI 416
 DB 752 VMSMQ-STLDSERIRSDALRV 773

RESULT 37

Q26080 PRELIMINARY; PRT; 1950 AA.
 AC Q26080;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE MYOSIN HEAVY CHAIN.
 OS Placopecten magellanicus (Sea scallop).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoida;
 OC Pectinoidea; Pectinidae; Placopecten.
 OX NCBI_Taxid=6577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ADDUCTOR MUSCLE;
 RA Perisart-Micale, C, Kalabokis V., Nyilray, L, Szent-Gyorgyi A.G.;
 RT "Sequence variations in the surface loop near the nucleotide binding
 RT site modulate the ATP turnover rates of molluscan myosins.";
 RL J. Muscle Res. Cell Motil. 0:0-0(1996).
 DR EMBL: U59295; AAB03661.1;
 DR HSSP: P24733; 1MDC.
 DR Interpro: IPR000048; IQ.
 DR Interpro: IPR001609; myosin_head.
 DR Interpro: IPR004009; Myosin_N.
 DR Interpro: IPR002928; Myosin_tail.
 DR Interpro: IPR002017; Spectrin.
 DR Pfam: PF00612; IQ; 2.
 DR Pfam: PF00663; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR ProDom: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 DR PROSITE: PS50096; IQ; 1.
 SQ SEQUENCE 1950 AA; 224184 MW; 6FE358802352C19A CRC64;

Query Match 7.5%; Score 179; DB 5; Length 1950;
 Best Local Similarity 21.9%; Pred. No. 0.11;

Matches 95; Conservative 70; Mismatches 134; Indels 134; Gaps 18;

QY 55 KTIITLKFDP--LAOEENVLAEFLKNELDVSK-----AQLSQ 92
 DB 905 EKLMQKADPESEIKLEELKLBEDASDLEGKKMECDNANLKIDGELSHLSQKSE 964
 QY 93 KDBEKRDQAIDTL-----RDITLERN-AVVEISQ-----NALNKA 128
 DB 965 OEAKHKDQ--ISTLQEMSGQDEHIGLKNKKKALEEANKKTSISQAEDEKCNHLNK- 1021
 QY 129 EMLCSTLKQMKFLQDQDETQKAREAHRLKCKMK-----TMEQITLLQSQRSEVEEM 183
 DB 1022 --EKAKLEQALDELEDELLEKEKRYGVFERAKSKVEQDLKSTQENVEDLERVKR-ELEEN 1078
 QY 184 IR-----DMGVGQSAVEQLAVYCVSLKKEYENLKEARKATGELADRLKKDLVSS 232
 DB 1079 VRKKEAEITTLNSKLEDEQNLVSOLOKRIKELQRIELELEAEERNANBSKVEKQAEI 1138
 QY 233 RSKLKTNLTELQDA-----KLELSAQKDLQSA-----DOETISLRKSSDP 274
 DB 1139 NRELEELGEIRDEAGGATSAQIELNKKRAELKIRRDLEASIQHQAQISAIIRKKQD- 1197
 QY 275 PGNLEPASATNEVSVRLVFSPPAVEMMRLHOPPGDEIDLTNTFDVNTPTOTSGSQ 334

DB 1198 -----AANEMADQV--DQLOKVKSKSEKENNKMESENNEDLQAGI-----Q 1235
 QY 335 HCLPRLKLERARSPMGNVLKVKHKVSKPESQSLSGQRCVGEDELDELACGAPLFLIRNAV 394
 DB 1236 HISKNGCSEKVKQFES-----QVSDLNARLE-DSORSINEL----- 1272
 QY 395 LGOKOPNRTTAES 407
 DB 1273 --QSKSRLOAEN 1283

RESULT 38

Q9BWC0 PRELIMINARY; PRT; 471 AA.
 AC Q9BWC0;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE HYPOTHETICAL 54.6 KDA PROTEIN (FRAGMENT).
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=EYE, RETINOBLASTOMA;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC000280; AAH00280.1;
 DR Interpro: IPR002928; Myosin_tail.
 DR Interpro: IPR000533; Tropomyosin.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00194; TROPOMYOSIN.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 471 AA; 54589 MW; 82F1E4B83BE2E135 CRC64;

Query Match 7.5%; Score 178.5; DB 4; Length 471;
 Best Local Similarity 22.0%; Pred. No. 0.027;
 Matches 75; Conservative 75; Mismatches 136; Indels 55; Gaps 10;

QY 75 DAEFLKNELDSVKAQSLSQDKREKRDQAIDTLRDITLEERNATVESIQNALNKAEMLCST 134
 DB 7 DMEQLSSKDDVGNVHELEKSKRALQGYEEMKQTQLELEDELQATVEDAKRLREVMQA 66
 QY 135 LKQMKFLQDQDE-----TKQAREAHRLKCKMK-----TMEQITLLQSORS 178
 DB 67 MKAQFERDLQTRDQNEKRRLLIKQVRELEAELEDEKORALAVASKKMEIDLKDLA 126
 QY 179 EVE--EMIRDMGVGQSAVEQLAVYCVSLKKEYENLKEARKATGELADRLKKDLVSSSKL 236
 DB 127 QIEANKARD-----EYIKDLRKLOAQMKDYORELEEARASRDEIFQSKK-----SEKTL 177
 QY 237 KTIITLQAKLLELSAQKDLQSADEITSLRKSSDPNLPASATNTVSR----- 291
 DB 178 KSLAEITLQLEELASSERARRHAQGERDEL---ADEITNSASGSKSLDEKRLERLARIA 234
 QY 292 -----VFESPAPVEMMRLHOPPGDEIDLTNTFDVNTPTOTSGQHLPK---KLC 342
 DB 235 QLEBELEEDQSNMELDLDRKK-----TTLQVDTNIAELAEASRAAOKSDNARQO 284
 QY 343 LERARSPMGNVLKVKHKVSKPESQSLSG--QRCVGEDEL 382
 DB 285 LERONKELKAKLOLEBAGVSKFRATISALEAKIGOLEOL 325

RESULT 39

Q63731 PRELIMINARY; PRT; 1999 AA.
 AC Q63731;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE NEURONAL MYOSIN HEAVY CHAIN.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN.
RX MEDLINE=92235856; PubMed=1569576;
RA Sun W., Chantler P.D.;
RT "Cloning of the cDNA encoding a neuronal myosin heavy chain from
RT mammalian brain and its differential expression within the central
RT nervous system[see comments]."
RL J. Mol. Biol. 224:1185-1193(1992).
DR HSP; P10587; 1BR2.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR001609; myosin_head.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 2.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KM Myosin.
SQ SEQUENCE 1999 AA; 231456 MW; 090C181D55727B41 CRC64;

Query Match 7.58; Score 178.5; DB 11; Length 1999;
Best Local Similarity 20.98; Pred. No. 0.12;
Matches 93; Conservative 81; Mismatches 153; Indels 117; Gaps 15;

QY 48 QCRVGVKTTINFLPDLAEEENV-----LDNEF--LKNELDSVKAQLSOK 93
DB 1072 ELKMOALAKE--EELQALAVVEEDAOKNALKKIRELSQISELDSESRASRKA 129
QY 94 DREKRSQAIDLTLDLTLEERNATVESLQNALNKAEMLSYLK-----OMKPLEQ 144
DB 1130 EKQRKDGEELALAKTELEDLTDTAQQELRSKREGEVNLKKTLEBEAKTHAQIDEM 1189
QY 145 RODETKQAREFAHL--KCKMKTME-----QIELLOS-----QRSEV 180
DB 1190 ROKHSQAVEELAEQLEDTKRKAVANLEKAKOTLENERGELANEVYLLQGRDSEHKRRKY 1249
QY 181 EEMIRDMKV-----GQSAVEQLAVYCVSLKREYENLKEKRRKATGELADRLKDYSSRKL 236
DB 1250 EAQLOELQVKNESGERRTYELADVTKLQVELDVTGLLSQSDSKSKSLTRDFSALESOL 1309
QY 237 K-----TLNTELOAKLELRSQKDLQSAQD-----ITSL 267
DB 1310 QDTQELQDEENKOKLSLSTKQVDEKNSFREQLSEEEBAKHLEKQITLNAQVADM 1369
QY 268 RKKSDDPPGNLEPASATN-----ETVSRVLVFSPA--PVEKMNPLHOPFGDEIDL 317
DB 1370 KKKKEDSVGCLETAEEVVRKLOKDLDEGLSHRHEEKVAAYDKLEKTKRLQDELDDLVDL 1429
QY 318 N-----TTPDVNTPPPTQSSQHCCLRKLCLEKARBSMOVNLKKVHKVSPE 364
DB 1430 DHQROSACNLEKKOKKFP-----QLAAEITTSKAKYAEERARDAEERAERAKTRELSLAR 1483
QY 365 SOLSIGQRC-----VGLDEEL 382
DB 1484 ALEMEKQAKFLRKNLOEMTERL 1507

RESULT 40
Q9VDZ1

ID Q9VDZ1 PRELIMINARY; PRT; 482 AA.
AC Q9VDZ1;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CG5555 PROTEIN.
GN Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRAIN-BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Morten J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Plannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Catheu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davaport L.B., Davies P.,
RA De Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ideyem C.,
RA Jajal M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshneft A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazotto M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1. SIMILARITY: CONFAINS 1 RING-TYPE ZINC FINGER.
DR EMBL: AF003725; AAF55646.1; -
DR FLYBase: FBgn0038686; CG5555.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001607; zf-UBP.
DR InterPro: IPR001841; znf_fing.
DR Pfam: PF00097; zf-C3HC4; 1.
DR Pfam: PF02148; zf-UBP; 1.
DR SMART: SM00184; RING; 1.
DR SMART: SM00290; znf-UBP; 1.
KW Zinc-finger.
SQ SEQUENCE 482 AA; 54620 MW; 89A9D9D45CDEF39F CRC64;

Query Match 7.48; Score 178; DB 5; Length 482;
Best Local Similarity 20.08; Pred. No. 0.03;
Matches 74; Conservative 56; Mismatches 96; Indels 144; Gaps 14;

QY 4 LSLCTGSDPFDHSRD-VAAIHGHTFHLOCLIQWFETAPSRTPCR----- 50

```

Db 172 LPTCPVCLERNDSEVDGLTILCNHAFHASCIMKMGDS-----TCPVCRHVQTPGLVEDSV 227
QY 51 -----IQVCKRTI-----INK 61
Db 228 CMECEGTSLMICLICGHVGGGRYOGHAAHFHATNHTFAMQIGTSSVWDYAGDNFVHR 287
QY 62 LFFDLAOREENVLDAEFLKNEIDSVKAOISQKREKRDQAIIIDTLRDITLREERNATVESL 121
Db 288 LF-----QKSDGKLV-----ASQTEKD-----EREKIDSM 314
QY 122 QNALNKAEMLCSTLKKQKFLBQRODETKQA---REBAHRLKCKMKTMEQITELLQSOR 177
Db 315 Q-MEFTYLTLSQLDTORRYEERMERLEOEQONHKATANDAKTEVSELOOQNMQKEX 372
QY 178 SEVEEMINDMGVGSVAEQAVYCVSLKKEYENLKEARKATGELADRLKDLVSSRSKLK 237
Db 373 VNLE-----RKLQHTAKLKDVOQOLNEER---ELSKALQSNQSSWHGKXK 415
QY 238 TLNTELDOAKLELRSQKDLQSADEITSLRKKSDD-----PGNLEPASATNETV 288
Db 416 LLEQOYNFEK-----QTHDAEYTELKQDLRDIMFPLDNQOKLANTEIAGS---TV 462
QY 289 SRLVFESPAP 298
Db 463 TGIKEKEPDP 472

```

Search completed: September 4, 2002, 16:15:58
 Job time: 1482 sec